STIC Search Report
Biotech-Chem Library

STIC Database Tracking Number: 139402

TO: Janet Epps-Ford
Location: REM-2C05/2C18
Art Unit: 1635
Thursday, December 02, 2004
Case Serial Number: 09/551494

From: Paul Schulwitz
Location: Biotech-Chem Library
REM-1A65
Phone: (571)272-2527

paul.schulwitz@uspto.gov

Search Notes

Examiner Epps-Ford,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Paul Schulwitz
Technical Information Specialist
STIC Biotech/Chem Library
(571)272-2527

Schulwitz, Paul

From: Epps-Ford, Janet
Sent: Wednesday, December 01, 2004 12:33 PM
To: Schulwitz, Paul
Subject: Sequence and Word-Search

Hi there, I was wondering if I can get the following claim language searched:

Application 09/551,494

An RNA viral vector comprising the nucleotide sequence of SEQ ID NO: 5 from the nucleotide at position 5430 to the nucleotide at position 5505.

Thanks,
Janet L. Epps-Ford, Ph.D.
Art Unit 1635
Mailbox: Remsen 2C18
Office: Remsen 2C05
Phone: 571-272-0757
Fax: 571-273-0757

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 1, 2004, 20:12:59 ; Search time 408 Seconds
(without alignments)
977.833 Million cell updates/sec

Title: US-09-551-494-5_COPY_5430_5505
Perfect score: 76
Sequence: 1 gtgacagacggctcgccaat.....tgaagtaccaatggctgtga 76

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	76	100.0	6355	3	AAC62372	Aac62372 cDNA sequ
	2	76	100.0	6355	9	ACC85005	Acc85005 TMV-U2 ge
	3	69.6	91.6	769	10	ADI26338	Adi26338 Novel end
	4	69.6	91.6	769	11	ADM68451	Adm68451 Tobacco m
	5	69.6	91.6	769	12	ADP26604	Adp26604 Tobamovir
c	6	68	89.5	772	10	ADI26344	Adi26344 Novel end
c	7	68	89.5	772	11	ADM68457	Adm68457 Mosaic vi
c	8	68	89.5	772	12	ADP26610	Adp26610 Heterodup
	9	58.6	77.1	411	3	AAC62379	Aac62379 Origin of
c	10	48.8	64.2	769	10	ADI26343	Adi26343 Novel end
c	11	48.8	64.2	769	11	ADM68456	Adm68456 Mosaic vi
c	12	48.8	64.2	769	12	ADP26609	Adp26609 Heterodup
	13	45.8	60.3	5997	2	AAQ12188	Aaq12188 Odontoglo
	14	45.8	60.3	6597	2	AAQ38106	Aaq38106 ORSV cDNA
	15	44	57.9	356	1	AAN30116	Aan30116 TMV-RNA f
	16	44	57.9	1019	1	AAN30115	Aan30115 TMV-RNA f
c	17	42.4	55.8	557	10	ADD17931	Add17931 DNA (SeqI
c	18	42.4	55.8	557	10	ADK56952	Adk56952 Plant DNA
	19	42.4	55.8	632	2	AAV16847	Aav16847 Tobacco m
c	20	42.4	55.8	792	10	ADI26342	Adi26342 Novel end
c	21	42.4	55.8	792	11	ADM68455	Adm68455 Mosaic vi
c	22	42.4	55.8	792	12	ADP26608	Adp26608 Heterodup
	23	42.4	55.8	804	12	ADN63297	Adn63297 Tobamovir
	24	42.4	55.8	805	6	ABT06582	Abt06582 Wild-type
	25	42.4	55.8	807	6	ABT06572	Abt06572 Wild-type
	26	42.4	55.8	807	8	ABX14577	Abx14577 Tomato mo
	27	42.4	55.8	807	11	ADM68440	Adm68440 Tobacco m
	28	42.4	55.8	807	12	ADP26593	Adp26593 Tobamovir
	29	42.4	55.8	895	2	AAQ62653	Aaq62653 Tobacco-m
	30	42.4	55.8	1187	10	ADD17932	Add17932 DNA (SeqI
	31	42.4	55.8	1187	10	ADK56953	Adk56953 Plant DNA
	32	42.4	55.8	1825	2	AAQ67663	Aaq67663 TMV repli
	33	42.4	55.8	1825	2	AAV54825	Aav54825 Replicon
	34	42.4	55.8	1971	3	AAC62382	Aac62382 cDNA sequ
	35	42.4	55.8	5484	8	ADA15011	Ada15011 Tobacco m
	36	42.4	55.8	6395	2	AAQ95155	Aaq95155 Tobacco m
	37	42.4	55.8	6395	2	AAZ20642	Aaz20642 TMV-based
	38	42.4	55.8	6395	3	AAC62369	Aac62369 cDNA sequ
	39	42.4	55.8	6395	4	AAF82330	Aaf82330 Wild-type
	40	42.4	55.8	6425	2	AAZ20644	Aaz20644 TMV-based
	41	42.4	55.8	6425	4	AAF82332	Aaf82332 Tobacco m
	42	42.4	55.8	6439	2	AAZ20643	Aaz20643 TMV-based
	43	42.4	55.8	6439	4	AAF82331	Aaf82331 Tobacco m
	44	42.4	55.8	6446	2	AAZ20645	Aaz20646 TMV-based
	45	42.4	55.8	6446	4	AAF82334	Aaf82334 Tobacco m

ALIGNMENTS

RESULT 1
AAC62372
ID AAC62372 standard; DNA; 6355 BP.
XX
AC AAC62372;
XX
DT 19-MAR-2001 (first entry)
XX
DE cDNA sequence of the genome of tobacco mosaic virus-U2.
XX
KW Plant phenotype; gene trait; Nicotiana; Oryza sativa; Zea mays; Brassica;
KW Gossypum; Triticum; Arabidopsis; Petunia; herbicide; transgenic plant;
KW tobacco necrosis virus; TNV; tobacco mosaic virus; TMV; helper virus; ss.
XX
OS Tobacco mosaic virus.
XX
PN WO200063397-A2.
XX
PD 26-OCT-2000.
XX
PF 17-APR-2000; 2000WO-EP003521.
XX
PR 20-APR-1999; 99US-00294022.
XX
PA (AVET) AVENTIS CROPSCIENCE NV.
XX
PI Meulewaeter F, Cornelissen M, Jacobs J, Van Eldik G, Metzlaff M;
XX
DR WPI; 2000-687182/67.
XX
PT Identifying and isolating genes involved in determining the trait or
PT phenotype of plant species, by infecting plants with gene silencing
PT constructs targeted to the gene, and identifying plants with altered
PT traits.
XX
PS Example 1; Page 53-56; 64pp; English.
XX
CC The specification describes a method for isolating genes that determine a
CC trait or phenotype of a plant species. The method comprises identifying a
CC set of nucleic acids of genes correlated with the trait, creating a
CC library of gene silencing constructs in a viral RNA vector, targeting the
CC gene silencing constructs to the nucleic acid set, infecting a collection
CC of individual plants with these, identifying plants with altered traits
CC or phenotype, and isolating genes of the invention. The method is useful

CC for isolating genes involved in the determination of trait or a phenotype
CC of a plant such as Nicotiana, Oryza sativa, Zea mays, Brassica, Gossypum,
CC Triticum, Arabidopsis or Petunia. The method is also useful for
CC modulating the expression of selected nucleic acid sequences and for
CC validating the function of a nucleic acid sequence whose expression is
CC correlated with the presence or absence of a specific trait in plants,
CC but with otherwise unknown function. The method is also useful for
CC developing agronomically useful products such herbicides or transgenic
CC plants. The present sequence represents the cDNA sequence of the genome
CC of tobacco mosaic virus (TMV)-U2. The sequence was used to construct a
CC plasmid vector for the synthesis of an infective hybrid tobacco mosaic
CC virus (TMV)/satellite tobacco necrosis virus (STNV) helper virus RNA.
CC This helper virus is used in the method of the invention
XX
SQ Sequence 6355 BP; 1933 A; 1112 C; 1489 G; 1821 T; 0 U; 0 Other;

Query Match 100.0%; Score 76; DB 3; Length 6355;
Best Local Similarity 100.0%; Pred. No. 2.2e-16;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5430 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 5489

Qy 61 GTACCAATGGCTGTGA 76
||||||||||||||||
Db 5490 GTACCAATGGCTGTGA 5505

RESULT 2
ACC85005
ID ACC85005 standard; DNA; 6355 BP.
XX
AC ACC85005;
XX
DT 13-OCT-2003 (first entry)
XX
DE TMV-U2 genome nucleotide sequence.
XX
KW Inhibitory RNA; viral RNA vector; coat protein; TMV; U2; gene; ds.
XX
OS Tobacco mosaic virus.
XX
PN WO2003052108-A2.
XX
PD 26-JUN-2003.
XX
PF 05-DEC-2002; 2002WO-EP013964.
XX
PR 18-DEC-2001; 2001US-0340488P.
XX
PA (FARB) BAYER BIOSCIENCE NV.
XX
PI Metzlaff MH, Gossele VML, Meulewaeter F, Fache ICA;
XX
DR WPI; 2003-523529/49.
XX
PT Introducing inhibitory RNA into a plant cell comprises providing a viral
PT RNA vector derived from a satellite RNA virus that encodes a coat
PT protein, and infecting a plant with the viral RNA vector and a
PT corresponding helper virus.
XX
PS Example; Page 79-82; 86pp; English.
XX
CC The invention relates to introducing inhibitory RNA into a plant cell.
CC The method involves providing a viral RNA vector derived from a satellite
CC RNA virus having a sequence that encodes a coat protein, and infecting a
CC plant with the viral RNA vector and a corresponding helper virus. The
CC methods and viral RNA vectors are useful in introducing inhibitory RNA
CC into plant cells. These may be used to determine or validate the function
CC of isolated nucleic acid sequences in plants. The present sequence
CC represents the nucleotide sequence of the genome of tobacco mosaic virus
CC (TMV)-U2
XX
SQ Sequence 6355 BP; 1933 A; 1112 C; 1489 G; 1821 T; 0 U; 0 Other;

Query Match 100.0%; Score 76; DB 9; Length 6355;
Best Local Similarity 100.0%; Pred. No. 2.2e-16;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5430 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 5489

Qy 61 GTACCAATGGCTGTGA 76
||||||||||||||||
Db 5490 GTACCAATGGCTGTGA 5505

RESULT 3
ADI26338
ID ADI26338 standard; DNA; 769 BP.
XX
AC ADI26338;
XX
DT 22-APR-2004 (first entry)
XX
DE Novel endonuclease Res I-related clone DNA 3.
XX
KW endonuclease; molecular biology; plant propagation; phenotypic trait;
KW herbicide tolerance; heat tolerance; cold tolerance; drought; salinity;
KW osmotic stress; pest resistance; insect; nematode; arachnid; fungal;
KW bacterial; viral; enzyme production; secondary metabolite;
KW male sterility; female sterility; dwarfness; early maturity; Res I; ds.
XX
OS Tobacco mosaic virus.
XX
PN US2003148315-A1.
XX
PD 07-AUG-2003.
XX
PF 01-AUG-2002; 2002US-00211079.
XX
PR 01-FEB-2002; 2002US-0353722P.
PR 14-MAR-2002; 2002US-00098155.
XX
PA (PADG/) PADGETT H S.
PA (VAEW/) VAEWHONGS A A.
XX
PI Padgett HS, Vaewhongs AA;
XX
DR WPI; 2003-897548/82.
XX
PT New nucleic acid molecule encoding endonucleases, useful in molecular
PT biology, specifically to generating populations of related nucleic acid
PT molecules, and in plant propagation with useful phenotypic traits.
XX
PS Example 15; Fig 8; 46pp; English.
XX
CC This invention relates to a novel endonuclease (Res I) nucleic acid
CC molecule which comprises a fully defined sequence of 899 bp given in the
CC specification. The methods and compositions of the present invention are
CC useful in molecular biology, and more specifically to generating
CC populations of related nucleic acid molecules. They may also be used in
CC plant propagation with useful phenotypic traits, such as improved
CC tolerance to herbicides, improved tolerance to extremes of heat or cold,
CC drought, salinity or osmotic stress, improved resistance to pests
CC (insects, nematodes or arachnids) or diseases (fungal, bacterial or
CC viral), production of enzymes or secondary metabolites, male or female
CC sterility, dwarfness and early maturity. The present sequence is that of
CC a clone which was derived during the exemplification of the invention.
XX
SQ Sequence 769 BP; 247 A; 102 C; 201 G; 219 T; 0 U; 0 Other;

Query Match 91.6%; Score 69.6; DB 10; Length 769;
Best Local Similarity 94.7%; Pred. No. 2.3e-14;

```
  Matches   72;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              || |||||||||||||||||||||||||||||||| ||||||||||||||| | ||||||
Db        541 GTAACAGACGGCTCGCCAATTGAACTCACTGAAAAAGTTGTTGAGGAGTTCATAGATGAA 600

Qy         61 GTACCAATGGCTGTGA 76
              ||||||||||||||||
Db        601 GTACCAATGGCTGTGA 616


RESULT 4
ADM68451
ID   ADM68451 standard; DNA; 769 BP.
XX
AC   ADM68451;
XX
DT   03-JUN-2004  (first entry)
XX
DE   Tobacco mosaic virus movement protein gene #4.
XX
KW   ds; mismatch endonuclease; endonuclease; gene shuffling technology;
KW   single nucleotide polymorphism; cancer susceptibility;
KW   sequence variation redistribution; movement protein; gene.
XX
OS   Tobacco mosaic virus.
XX
PN   US2003157682-A1.
XX
PD   21-AUG-2003.
XX
PF   31-JAN-2003; 2003US-00356708.
XX
PR   01-FEB-2002; 2002US-0353722P.
PR   14-MAR-2002; 2002US-00098155.
PR   01-AUG-2002; 2002US-00211079.
XX
PA   (PADG/) PADGETT H S.
PA   (VAEW/) VAEWHONGS A A.
PA   (VOJD/) VOJDANI F S.
PA   (SMIT/) SMITH M L.
PA   (LIND/) LINDBO J A.
PA   (FITZ/) FITZMAURICE W P.
XX
PI   Padgett HS,  Vaewhongs AA,  Vojdani FS,  Smith ML,  Lindbo JA;
PI   Fitzmaurice WP;
XX
DR   WPI; 2003-766176/72.
XX
PT   Making a mismatch endonuclease, useful in gene shuffling and in detection
PT   of single nucleotide polymorphisms, comprises transfecting a host with a
PT   recombinant viral vector including a polynucleotide encoding a mismatch
PT   endonuclease.
XX
PS   Example 14; SEQ ID NO 20; 79pp; English.
XX
CC   The invention relates to a method of making a mismatch endonuclease
CC   enzyme comprising transfecting a host plant, animal, yeast, fungus or
CC   bacterium with a recombinant viral vector that encodes a polynucleotide
CC   sequence for a mismatch endonuclease, growing the host so that the
CC   polynucleotide is expressed, and extracting the mismatch endonuclease
CC   enzyme from the host. The method is useful for making mismatch
CC   endonuclease enzymes, for obtaining peptides and polynucleotides with
CC   desired functional properties and for detecting mutations. The mismatch
CC   endonuclease enzymes are useful in gene shuffling technology for
CC   developing new genes, in detecting single nucleotide polymorphisms for
CC   e.g. detecting evidence of cancer susceptibility, or in redistributing
CC   sequence variations between non-identical polynucleotide sequences. The
CC   present sequence represents a tobacco mosaic virus movement protein gene.
XX
SQ   Sequence 769 BP; 247 A; 102 C; 201 G; 219 T; 0 U; 0 Other;

  Query Match             91.6%;  Score 69.6;  DB 11;  Length 769;

  Best Local Similarity   94.7%;  Pred. No. 2.3e-14;
  Matches   72;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              || |||||||||||||||||||||||||||||||| ||||||||||||||| | ||||||
Db        541 GTAACAGACGGCTCGCCAATTGAACTCACTGAAAAAGTTGTTGAGGAGTTCATAGATGAA 600

Qy         61 GTACCAATGGCTGTGA 76
              ||||||||||||||||
Db        601 GTACCAATGGCTGTGA 616


RESULT 5
ADP26604
ID   ADP26604 standard; DNA; 769 BP.
XX
AC   ADP26604;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Tobamovirus DNA #5.
XX
KW   Sequence variation; heteroduplex; transcription; DNA integration;
KW   ribozyme expression; gene; ds; tobamovirus.
XX
OS   Tobamovirus.
XX
PN   US2004110130-A1.
XX
PD   10-JUN-2004.
XX
PF   25-OCT-2002; 2002US-00280913.
XX
PR   02-FEB-2001; 2001US-0266386P.
PR   14-FEB-2001; 2001US-0268785P.
PR   01-FEB-2002; 2002US-00066398.
PR   08-AUG-2002; 2002US-0402342P.
XX
PA   (LARG-) LARGE SCALE BIOLOGY CORP.
XX
PI   Padgett HS,  Lindbo JA,  Fitzmaurice WP;
XX
DR   WPI; 2004-440326/41.
XX
PT   Redistributing sequence variations between non-identical polynucleotide
PT   sequences, useful for generating improved polynucleotide having a desired
PT   characteristic, comprises making a heteroduplex and introducing a nick.
XX
PS   Example 15; SEQ ID NO 20; 75pp; English.
XX
CC   The invention relates to an in vitro method of redistributing sequence
CC   variations between non-identical polynucleotide sequences, comprising
CC   making a heteroduplex polynucleotide from two non-identical
CC   polynucleotides, introducing a nick in the second strand at or near a
CC   base pair mismatch site, removing the mismatched base(s) from the
CC   mismatch site where the nick occurred and using the first strand as a
CC   template to replace the removed base(s) with bases that complement the
CC   base(s) in the first strand. The invention also relates to an in vitro
CC   method of making a population of sequence variants from a heteroduplex
CC   polynucleotide sequence, obtaining a polynucleotide sequence encoding a
CC   desired functional property and identifying a reasserted DNA molecule
CC   encoding a protein with a desired functional property. The method is
CC   useful for generating an improved polynucleotide sequence or a population
CC   of improved polynucleotide sequences possessing at least one desired
CC   phenotypic characteristic (e.g., promotes transcription of linked
CC   polynucleotides), where such polynucleotides are useful for expression
CC   from a plant, animal, fungal, yeast, or bacterial expression vector, for
CC   integration to form a transgenic plant, animal or microorganism, and for
CC   expression of a ribozyme. This sequence represents DNA used in the scope
CC   of the invention.
XX
SQ   Sequence 769 BP; 247 A; 102 C; 201 G; 219 T; 0 U; 0 Other;
```

```
  Query Match             91.6%;  Score 69.6;  DB 12;  Length 769;
  Best Local Similarity   94.7%;  Pred. No. 2.3e-14;
  Matches   72;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              || |||||||||||||||||||||||||||||| ||||||||||||||| | ||||||
Db        541 GTAACAGACGGCTCGCCAATTGAACTCACTGAAAAAGTTGTTGAGGAGTTCATAGATGAA 600

Qy         61 GTACCAATGGCTGTGA 76
              ||||||||||||||||
Db        601 GTACCAATGGCTGTGA 616


RESULT 6
ADI26344/c
ID   ADI26344 standard; DNA; 772 BP.
XX
AC   ADI26344;
XX
DT   22-APR-2004  (first entry)
XX
DE   Novel endonuclease Res I-related clone DNA 9.
XX
KW   endonuclease; molecular biology; plant propagation; phenotypic trait;
KW   herbicide tolerance; heat tolerance; cold tolerance; drought; salinity;
KW   osmotic stress; pest resistance; insect; nematode; arachnid; fungal;
KW   bacterial; viral; enzyme production; secondary metabolite;
KW   male sterility; female sterility; dwarfness; early maturity; Res I; ds.
XX
OS   Tobacco mosaic virus.
OS   Tomato mosaic virus.
XX
PN   US2003148315-A1.
XX
PD   07-AUG-2003.
XX
PF   01-AUG-2002; 2002US-00211079.
XX
PR   01-FEB-2002; 2002US-0353722P.
PR   14-MAR-2002; 2002US-00098155.
XX
PA   (PADG/) PADGETT H S.
PA   (VAEW/) VAEWHONGS A A.
XX
PI   Padgett HS,  Vaewhongs AA;
XX
DR   WPI; 2003-897548/82.
XX
PT   New nucleic acid molecule encoding endonucleases, useful in molecular
PT   biology, specifically to generating populations of related nucleic acid
PT   molecules, and in plant propagation with useful phenotypic traits.
XX
PS   Example 15; Fig 14; 46pp; English.
XX
CC   This invention relates to a novel endonuclease (Res I) nucleic acid
CC   molecule which comprises a fully defined sequence of 899 bp given in the
CC   specification. The methods and compositions of the present invention are
CC   useful in molecular biology, and more specifically to generating
CC   populations of related nucleic acid molecules. They may also be used in
CC   plant propagation with useful phenotypic traits, such as improved
CC   tolerance to herbicides, improved tolerance to extremes of heat or cold,
CC   drought, salinity or osmotic stress, improved resistance to pests
CC   (insects, nematodes or arachnids) or diseases (fungal, bacterial or
CC   viral), production of enzymes or secondary metabolites, male or female
CC   sterility, dwarfness and early maturity. The present sequence is that of
CC   a clone which was derived during the exemplification of the invention.
XX
SQ   Sequence 772 BP; 227 A; 196 C; 109 G; 240 T; 0 U; 0 Other;

  Query Match             89.5%;  Score 68;  DB 10;  Length 772;
  Best Local Similarity   93.4%;  Pred. No. 8.2e-14;
  Matches   71;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              || |||||||||||||||||||||||||||||| ||||||| |||||| | ||||||
Db        232 GTAACAGACGGCTCGCCAATTGAACTCACTGAAAAAGTTGTTGATGAGTTCATAGATGAA 173

Qy         61 GTACCAATGGCTGTGA 76
              ||||||||||||||||
Db        172 GTACCAATGGCTGTGA 157


RESULT 7
ADM68457/c
ID   ADM68457 standard; DNA; 772 BP.
XX
AC   ADM68457;
XX
DT   03-JUN-2004  (first entry)
XX
DE   Mosaic virus movement protein gene GRAMMR clone #11.
XX
KW   ds; mismatch endonuclease; endonuclease; gene shuffling technology;
KW   single nucleotide polymorphism; cancer susceptibility;
KW   sequence variation redistribution; movement protein; gene.
XX
OS   Tobacco mosaic virus.
OS   Tomato mosaic virus.
XX
PN   US2003157682-A1.
XX
PD   21-AUG-2003.
XX
PF   31-JAN-2003; 2003US-00356708.
XX
PR   01-FEB-2002; 2002US-0353722P.
PR   14-MAR-2002; 2002US-00098155.
PR   01-AUG-2002; 2002US-00211079.
XX
PA   (PADG/) PADGETT H S.
PA   (VAEW/) VAEWHONGS A A.
PA   (VOJD/) VOJDANI F S.
PA   (SMIT/) SMITH M L.
PA   (LIND/) LINDBO J A.
PA   (FITZ/) FITZMAURICE W P.
XX
PI   Padgett HS,  Vaewhongs AA,  Vojdani FS,  Smith ML,  Lindbo JA;
PI   Fitzmaurice WP;
XX
DR   WPI; 2003-766176/72.
XX
PT   Making a mismatch endonuclease, useful in gene shuffling and in detection
PT   of single nucleotide polymorphisms, comprises transfecting a host with a
PT   recombinant viral vector including a polynucleotide encoding a mismatch
PT   endonuclease.
XX
PS   Example 14; SEQ ID NO 26; 79pp; English.
XX
CC   The invention relates to a method of making a mismatch endonuclease
CC   enzyme comprising transfecting a host plant, animal, yeast, fungus or
CC   bacterium with a recombinant viral vector that encodes a polynucleotide
CC   sequence for a mismatch endonuclease, growing the host so that the
CC   polynucleotide is expressed, and extracting the mismatch endonuclease
CC   enzyme from the host. The method is useful for making mismatch
CC   endonuclease enzymes, for obtaining peptides and polynucleotides with
CC   desired functional properties and for detecting mutations. The mismatch
CC   endonuclease enzymes are useful in gene shuffling technology for
CC   developing new genes, in detecting single nucleotide polymorphisms for
CC   e.g. detecting evidence of cancer susceptibility, or in redistributing
CC   sequence variations between non-identical polynucleotide sequences. The
CC   present sequence represents a mosaic virus movement protein gene GRAMMR
CC   clone.
XX
SQ   Sequence 772 BP; 227 A; 196 C; 109 G; 240 T; 0 U; 0 Other;

  Query Match             89.5%;  Score 68;  DB 11;  Length 772;
```

```
  Best Local Similarity   93.4%;  Pred. No. 8.2e-14;
  Matches   71;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              || |||||||||||||||||||||||||||||||| |||||||| |||||| | ||||||
Db        232 GTAACAGACGGCTCGCCAATTGAACTCACTGAAAAAGTTGTTGATGAGTTCATAGATGAA 173

Qy         61 GTACCAATGGCTGTGA 76
              ||||||||||||||||
Db        172 GTACCAATGGCTGTGA 157
```

RESULT 8
ADP26610/c
ID ADP26610 standard; DNA; 772 BP.
XX
AC ADP26610;
XX
DT 26-AUG-2004 (first entry)
XX
DE Heteroduplex DNA #15.
XX
KW Sequence variation; heteroduplex; transcription; DNA integration;
KW ribozyme expression; gene; ds.
XX
OS Synthetic.
XX
PN US2004110130-A1.
XX
PD 10-JUN-2004.
XX
PF 25-OCT-2002; 2002US-00280913.
XX
PR 02-FEB-2001; 2001US-0266386P.
PR 14-FEB-2001; 2001US-0268785P.
PR 01-FEB-2002; 2002US-00066390.
PR 08-AUG-2002; 2002US-0402342P.
XX
PA (LARG-) LARGE SCALE BIOLOGY CORP.
XX
PI Padgett HS, Lindbo JA, Fitzmaurice WP;
XX
DR WPI; 2004-440326/41.
XX
PT Redistributing sequence variations between non-identical polynucleotide
PT sequences, useful for generating improved polynucleotide having a desired
PT characteristic, comprises making a heteroduplex and introducing a nick.
XX
PS Example 15; SEQ ID NO 26; 75pp; English.
XX
CC The invention relates to an in vitro method of redistributing sequence
CC variations between non-identical polynucleotide sequences, comprising
CC making a heteroduplex polynucleotide from two non-identical
CC polynucleotides, introducing a nick in the second strand at or near a
CC base pair mismatch site, removing the mismatched base(s) from the
CC mismatch site where the nick occurred and using the first strand as a
CC template to replace the removed base(s) with bases that complement the
CC base(s) in the first strand. The invention also relates to an in vitro
CC method of making a population of sequence variants from a heteroduplex
CC polynucleotide sequence, obtaining a polynucleotide sequence encoding a
CC desired functional property and identifying a reasserted DNA molecule
CC encoding a protein with a desired functional property. The method is
CC useful for generating an improved polynucleotide sequence or a population
CC of improved polynucleotide sequences possessing at least one desired
CC phenotypic characteristic (e.g., promotes transcription of linked
CC polynucleotides), where such polynucleotides are useful for expression
CC from a plant, animal, fungal, yeast, or bacterial expression vector, for
CC integration to form a transgenic plant, animal or microorganism, and for
CC expression of a ribozyme. This sequence represents DNA used in the scope
CC of the invention.
XX
SQ Sequence 772 BP; 227 A; 196 C; 109 G; 240 T; 0 U; 0 Other;

```
  Query Match             89.5%;  Score 68;  DB 12;  Length 772;
  Best Local Similarity   93.4%;  Pred. No. 8.2e-14;
  Matches   71;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              || |||||||||||||||||||||||||||||||| |||||||| |||||| | ||||||
Db        232 GTAACAGACGGCTCGCCAATTGAACTCACTGAAAAAGTTGTTGATGAGTTCATAGATGAA 173

Qy         61 GTACCAATGGCTGTGA 76
              ||||||||||||||||
Db        172 GTACCAATGGCTGTGA 157
```

RESULT 9
AAC62379
ID AAC62379 standard; DNA; 411 BP.
XX
AC AAC62379;
XX
DT 19-MAR-2001 (first entry)
XX
DE Origin of assembly (OAS) of a tobacco mosaic virus (TMV)-U2.
XX
KW Plant phenotype; gene trait; Nicotiana; Oryza sativa; Zea mays; Brassica;
KW Gossypum; Triticum; Arabidopsis; Petunia; herbicide; transgenic plant;
KW tobacco necrosis virus; TNV; tobacco mosaic virus; TMV; helper virus;
KW origin of assembly; ss.
XX
OS Tobacco mosaic virus.
XX
PN WO200063397-A2.
XX
PD 26-OCT-2000.
XX
PF 17-APR-2000; 2000WO-EP003521.
XX
PR 20-APR-1999; 99US-00294022.
XX
PA (AVET) AVENTIS CROPSCIENCE NV.
XX
PI Meulewaeter F, Cornelissen M, Jacobs J, Van Eldik G, Metzlaff M;
XX
DR WPI; 2000-687182/67.
XX
PT Identifying and isolating genes involved in determining the trait or
PT phenotype of plant species, by infecting plants with gene silencing
PT constructs targeted to the gene, and identifying plants with altered
PT traits.
XX
PS Example 1; Page 63; 64pp; English.
XX
CC The specification describes a method for isolating genes that determine a
CC trait or phenotype of a plant species. The method comprises identifying a
CC set of nucleic acids of genes correlated with the trait, creating a
CC library of gene silencing constructs in a viral RNA vector, targeting the
CC gene silencing constructs to the nucleic acid set, infecting a collection
CC of individual plants with these, identifying plants with altered traits
CC or phenotype, and isolating genes of the invention. The method is useful
CC for isolating genes involved in the determination of trait or a phenotype
CC of a plant such as Nicotiana, Oryza sativa, Zea mays, Brassica, Gossypum,
CC Triticum, Arabidopsis or Petunia. The method is also useful for
CC modulating the expression of selected nucleic acid sequences and for
CC validating the function of a nucleic acid sequence whose expression is
CC correlated with the presence or absence of a specific trait in plants,
CC but with otherwise unknown function. The method is also useful for
CC developing agronomically useful products such herbicides or transgenic
CC plants. The present sequence is an origin of assembly (OAS) of a tobacco
CC mosaic virus (TMV)-U2. The sequence is used to construct infective hybrid
CC tobacco mosaic virus (TMV)/tobacco necrosis virus (TNV) vectors, for use
CC in the method of the invention
XX
SQ Sequence 411 BP; 140 A; 70 C; 93 G; 108 T; 0 U; 0 Other;

```
  Query Match             77.1%;  Score 58.6;  DB 3;  Length 411;
  Best Local Similarity   93.8%;  Pred. No. 1.4e-10;
  Matches   61;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         12 CTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAAGTACCAATGGC 71
              ||||||||||||||||||||||||| |||||||| |||||||| |||||||||| |||||
Db          3 CTCGCCAATTGAACTCACTGAAAAAGTTGTTGATGAGTTCGTAGATGAAGTACCGATGGC 62

Qy         72 TGTGA 76
              |||||
Db         63 TGTGA 67


RESULT 10
ADI26343/c
ID   ADI26343 standard; DNA; 769 BP.
XX
AC   ADI26343;
XX
DT   22-APR-2004  (first entry)
XX
DE   Novel endonuclease Res I-related clone DNA 8.
XX
KW   endonuclease; molecular biology; plant propagation; phenotypic trait;
KW   herbicide tolerance; heat tolerance; cold tolerance; drought; salinity;
KW   osmotic stress; pest resistance; insect; nematode; arachnid; fungal;
KW   bacterial; viral; enzyme production; secondary metabolite;
KW   male sterility; female sterility; dwarfness; early maturity; Res I; ds.
XX
OS   Tobacco mosaic virus.
OS   Tomato mosaic virus.
XX
PN   US2003148315-A1.
XX
PD   07-AUG-2003.
XX
PF   01-AUG-2002; 2002US-00211079.
XX
PR   01-FEB-2002; 2002US-0353722P.
PR   14-MAR-2002; 2002US-00098155.
XX
PA   (PADG/) PADGETT H S.
PA   (VAEW/) VAEWHONGS A A.
XX
PI   Padgett HS,  Vaewhongs AA;
XX
DR   WPI; 2003-897548/82.
XX
PT   New nucleic acid molecule encoding endonucleases, useful in molecular
PT   biology, specifically to generating populations of related nucleic acid
PT   molecules, and in plant propagation with useful phenotypic traits.
XX
PS   Example 15; Fig 13; 46pp; English.
XX
CC   This invention relates to a novel endonuclease (Res I) nucleic acid
CC   molecule which comprises a fully defined sequence of 899 bp given in the
CC   specification. The methods and compositions of the present invention are
CC   useful in molecular biology, and more specifically to generating
CC   populations of related nucleic acid molecules. They may also be used in
CC   plant propagation with useful phenotypic traits, such as improved
CC   tolerance to herbicides, improved tolerance to extremes of heat or cold,
CC   drought, salinity or osmotic stress, improved resistance to pests
CC   (insects, nematodes or arachnids) or diseases (fungal, bacterial or
CC   viral), production of enzymes or secondary metabolites, male or female
CC   sterility, dwarfness and early maturity. The present sequence is that of
CC   a clone which was derived during the exemplification of the invention.
XX
SQ   Sequence 769 BP; 222 A; 205 C; 96 G; 246 T; 0 U; 0 Other;

  Query Match             64.2%;  Score 48.8;  DB 10;  Length 769;
  Best Local Similarity   77.6%;  Pred. No. 4.6e-07;
  Matches   59;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              |||| || ||    || || ||||| | ||| | |||||||| |||||| | ||||||
Db        229 GTGAACGATGGAGGACCCATGGAACTTTCAGAAGAAGTTGTTGATGAGTTCATAGATGAA 170

Qy         61 GTACCAATGGCTGTGA 76
              ||||||||||||||||
Db        169 GTACCAATGGCTGTGA 154


RESULT 11
ADM68456/c
ID   ADM68456 standard; DNA; 769 BP.
XX
AC   ADM68456;
XX
DT   03-JUN-2004  (first entry)
XX
DE   Mosaic virus movement protein gene GRAMMR clone #10.
XX
KW   ds; mismatch endonuclease; endonuclease; gene shuffling technology;
KW   single nucleotide polymorphism; cancer susceptibility;
KW   sequence variation redistribution; movement protein; gene.
XX
OS   Tobacco mosaic virus.
OS   Tomato mosaic virus.
XX
PN   US2003157682-A1.
XX
PD   21-AUG-2003.
XX
PF   31-JAN-2003; 2003US-00356708.
XX
PR   01-FEB-2002; 2002US-0353722P.
PR   14-MAR-2002; 2002US-00098155.
PR   01-AUG-2002; 2002US-00211079.
XX
PA   (PADG/) PADGETT H S.
PA   (VAEW/) VAEWHONGS A A.
PA   (VOJD/) VOJDANI F S.
PA   (SMIT/) SMITH M L.
PA   (LIND/) LINDBO J A.
PA   (FITZ/) FITZMAURICE W P.
XX
PI   Padgett HS,  Vaewhongs AA,  Vojdani FS,  Smith ML,  Lindbo JA;
PI   Fitzmaurice WP;
XX
DR   WPI; 2003-766176/72.
XX
PT   Making a mismatch endonuclease, useful in gene shuffling and in detection
PT   of single nucleotide polymorphisms, comprises transfecting a host with a
PT   recombinant viral vector including a polynucleotide encoding a mismatch
PT   endonuclease.
XX
PS   Example 14; SEQ ID NO 25; 79pp; English.
XX
CC   The invention relates to a method of making a mismatch endonuclease
CC   enzyme comprising transfecting a host plant, animal, yeast, fungus or
CC   bacterium with a recombinant viral vector that encodes a polynucleotide
CC   sequence for a mismatch endonuclease, growing the host so that the
CC   polynucleotide is expressed, and extracting the mismatch endonuclease
CC   enzyme from the host. The method is useful for making mismatch
CC   endonuclease enzymes, for obtaining peptides and polynucleotides with
CC   desired functional properties and for detecting mutations. The mismatch
CC   endonuclease enzymes are useful in gene shuffling technology for
CC   developing new genes, in detecting single nucleotide polymorphisms for
CC   e.g. detecting evidence of cancer susceptibility, or in redistributing
CC   sequence variations between non-identical polynucleotide sequences. The
CC   present sequence represents a mosaic virus movement protein gene GRAMMR
CC   clone.
XX
SQ   Sequence 769 BP; 222 A; 205 C; 96 G; 246 T; 0 U; 0 Other;

  Query Match             64.2%;  Score 48.8;  DB 11;  Length 769;
```

```
  Best Local Similarity   77.6%;  Pred. No. 4.6e-07;
  Matches   59;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              ||||  || ||    || || |||||  | ||| | |||||||| |||||| | ||||||
Db        229 GTGAACGATGGAGGACCCATGGAACTTTCAGAAGAAGTTGTTGATGAGTTCATAGATGAA 170

Qy         61 GTACCAATGGCTGTGA 76
              ||||||||||||||||
Db        169 GTACCAATGGCTGTGA 154


RESULT 12
ADP26609/c
ID   ADP26609 standard; DNA; 769 BP.
XX
AC   ADP26609;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Heteroduplex DNA #14.
XX
KW   Sequence variation; heteroduplex; transcription; DNA integration;
KW   ribozyme expression; gene; ds.
XX
OS   Synthetic.
XX
PN   US2004110130-A1.
XX
PD   10-JUN-2004.
XX
PF   25-OCT-2002; 2002US-00280913.
XX
PR   02-FEB-2001; 2001US-0266386P.
PR   14-FEB-2001; 2001US-0268785P.
PR   01-FEB-2002; 2002US-00056390.
PR   08-AUG-2002; 2002US-0402342P.
XX
PA   (LARG-) LARGE SCALE BIOLOGY CORP.
XX
PI   Padgett HS,  Lindbo JA,  Fitzmaurice WP;
XX
DR   WPI; 2004-440326/41.
XX
PT   Redistributing sequence variations between non-identical polynucleotide
PT   sequences, useful for generating improved polynucleotide having a desired
PT   characteristic, comprises making a heteroduplex and introducing a nick.
XX
PS   Example 15; SEQ ID NO 25; 75pp; English.
XX
CC   The invention relates to an in vitro method of redistributing sequence
CC   variations between non-identical polynucleotide sequences, comprising
CC   making a heteroduplex polynucleotide from two non-identical
CC   polynucleotides, introducing a nick in the second strand at or near a
CC   base pair mismatch site, removing the mismatched base(s) from the
CC   mismatch site where the nick occurred and using the first strand as a
CC   template to replace the removed base(s) with bases that complement the
CC   base(s) in the first strand. The invention also relates to an in vitro
CC   method of making a population of sequence variants from a heteroduplex
CC   polynucleotide sequence, obtaining a polynucleotide sequence encoding a
CC   desired functional property and identifying a reasserted DNA molecule
CC   encoding a protein with a desired functional property. The method is
CC   useful for generating an improved polynucleotide sequence or a population
CC   of improved polynucleotide sequences possessing at least one desired
CC   phenotypic characteristic (e.g., promotes transcription of linked
CC   polynucleotides), where such polynucleotides are useful for expression
CC   from a plant, animal, fungal, yeast, or bacterial expression vector, for
CC   integration to form a transgenic plant, animal or microorganism, and for
CC   expression of a ribozyme. This sequence represents DNA used in the scope
CC   of the invention.
XX
SQ   Sequence 769 BP; 222 A; 205 C; 96 G; 246 T; 0 U; 0 Other;

  Query Match             64.2%;  Score 48.8;  DB 12;  Length 769;
  Best Local Similarity   77.6%;  Pred. No. 4.6e-07;
  Matches   59;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              ||||  || ||    || || |||||  | ||| | |||||||| |||||| | ||||||
Db        229 GTGAACGATGGAGGACCCATGGAACTTTCAGAAGAAGTTGTTGATGAGTTCATAGATGAA 170

Qy         61 GTACCAATGGCTGTGA 76
              ||||||||||||||||
Db        169 GTACCAATGGCTGTGA 154


RESULT 13
AAQ12188
ID   AAQ12188 standard; DNA; 5997 BP.
XX
AC   AAQ12188;
XX
DT   27-AUG-2003  (revised)
DT   10-SEP-1991  (first entry)
XX
DE   Odontoglossum ring spot virus.
XX
KW   Odontoglossum ring spot virus; probe; coat protein; ORSV; ds.
XX
OS   Odontoglossum ringspot virus.
XX
PN   WO9108296-A.
XX
PD   13-JUN-1991.
XX
PF   28-NOV-1989;   89JP-00306626.
XX
PR   28-NOV-1989;   89JP-00306626.
XX
PA   (NIOC ) NIPPON OIL KK.
XX
PI   Isomura K,  Matsumoto Y,  Chatani M,  Ikegami M;
XX
DR   WPI; 1991-193200/26.
XX
PT   DNA obtd. by cleavage of cDNA corresp. to RNA - of Odontoglossum ring
PT   spot virus coding for viral coat protein and is useful as probe and
PT   vector for plant gene recombination.
XX
PS   Claim 2; Page 52-64; 84pp; Japanese.
XX
CC   The ds DNA is obtd. by cleavage with EcoRI of the cDNA corresp. to the
CC   genomic DNA. Organisms transformed with a vector contg. the DNA produce a
CC   peptide of mol.wt. of 33 kD, corresp. to the viral coat protein. The
CC   sequence may be used as probe, and for the prodn. of vectors for
CC   expression of plant genes. (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ   Sequence 5997 BP; 1787 A; 1032 C; 1311 G; 1867 T; 0 U; 0 Other;

  Query Match             60.3%;  Score 45.8;  DB 2;  Length 5997;
  Best Local Similarity   76.7%;  Pred. No. 8.9e-06;
  Matches   56;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              || |||| ||   ||| |  ||||| ||||||   |||||||| ||||||||||  ||
Db       4805 GTAACAGAAGGAGGGCCCACCGAACTTACTGAAGCAGTTGTTGATGAGTTCGTGGAGAAA 4864

Qy         61 GTACCAATGGCTG 73
              || || |||||||
Db       4865 GTTCCTATGGCTG 4877


RESULT 14
AAQ38106
ID   AAQ38106 standard; cDNA to mRNA; 6597 BP.
XX
```

```
AC   AAQ38106;
XX
DT   05-JUL-1993  (first entry)
XX
DE   ORSV cDNA.
XX
KW   Odontoglossum ring spot virus; screen; transformation; ds.
XX
OS   Odontoglossum ring spot virus.
XX
PN   JP05030975-A.
XX
PD   09-FEB-1993.
XX
PF   26-JUL-1991;   91JP-00276075.
XX
PR   26-JUL-1991;   91JP-00276075.
XX
PA   (NIOC ) NIPPON OIL KK.
XX
DR   WPI; 1993-087957/11.
XX
PT   cDNA of odontoglossum ring-spot virus gene - useful as vector for plants
PT   and probes for screening the virus.
XX
PS   Claim 2; Page 11-28; 30pp; Japanese.
XX
CC   RNA was isolated from Odontoglossum ring spot virus and used to prepare
CC   cDNA. The DNA or its restiction fragments can be used to screen for ORSV
CC   or to detect genes related to ORSV. Vectors contg. the DNA sequence can
CC   be used to transform E. coli, Bacillus subtilis, Agrobacterium or plant
CC   cells for prodn. of the recombinant 130K, 180K and 30K proteins of ORSV,
CC   coat proteins or their peptide fragments
XX
SQ   Sequence 6597 BP; 1970 A; 1170 C; 1425 G; 2032 T; 0 U; 0 Other;

  Query Match             60.3%;  Score 45.8;  DB 2;  Length 6597;
  Best Local Similarity   76.7%;  Pred. No. 9.2e-06;
  Matches   56;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              || ||||| ||   ||| |  ||||| ||||||   |||||||| ||||||||||  ||
Db       5405 GTAACAGAAGGAGGGCCCACCGAACTTACTGAAGCAGTTGTTGATGAGTTCGTGGAGAAA 5464

Qy         61 GTACCAATGGCTG 73
              || || |||||||
Db       5465 GTTCCTATGGCTG 5477


RESULT 15
AAN30116
ID   AAN30116 standard; RNA; 356 BP.
XX
AC   AAN30116;
XX
DT   25-MAR-2003  (revised)
DT   02-NOV-1992  (first entry)
XX
DE   TMV-RNA fragment I originating at the capped 5' end of the viral RNA and
DE   extending into the coat protein gene.
XX
KW   RNA plant virus vector; tobacco mosaic virus; ss.
XX
OS   Tobacco mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    5
FT                   /*tag=  e
FT                   /label= bp No. 5400
FT   misc_feature    61. .235
FT                   /*tag=  a
FT                   /label= nucleation region
FT   CDS             61. .235
FT                   /*tag=  c
FT                   /label= coat protein gene
FT   misc_signal     236. .320
FT                   /*tag=  b
FT                   /label= control region
FT   misc_feature    348. .356
FT                   /*tag=  d
FT                   /note= "site of cleavage by ribonuclease H"
XX
PN   EP67553-A.
XX
PD   22-DEC-1982.
XX
PF   27-MAY-1981;   81US-00267539.
XX
PR   27-MAY-1981;   81US-00267539.
XX
PA   (CANA ) NAT RES COUNCIL CANADA.
XX
PI   Pelcher LE,  Halasa MC;
XX
DR   WPI; 1983-00323K/01.
XX
PT   RNA plant virus vector from tobacco mosaic virus etc. - for modifying
PT   genes in plants to alter growth disease resistance etc.
XX
PS   Example; Page 27; 56pp; English.
XX
CC   The inventors claim an RNA plant virus vector from tocacco mosaic virus.
CC   The vector comprises a nucleotide sequence originating from the 5' end of
CC   the (+) strand of the viral RNA (Fragment I) (see AAN30116) and a
CC   sequence originating from the 3' end of the (+)strand (Fragment II) (see
CC   AAN30114, AAN30115). The RNA vector may have foreign genetic information
CC   inserted or attached, ultimately in the form of RNA, to the vector. The
CC   mode of Fragment I and II production fragments enables fragments of any
CC   desired length from any location in TMV-RNA to be generated (see
CC   AAN30117, AAN30118). Fragment I 76 is designated Frag.I CPG. The viral
CC   replicase gene is likely included Fragment I. (Updated on 25-MAR-2003 to
CC   correct PA field.)
XX
SQ   Sequence 356 BP; 121 A; 48 C; 87 G; 0 T; 100 U; 0 Other;

  Query Match             57.9%;  Score 44;  DB 1;  Length 356;
  Best Local Similarity   56.6%;  Pred. No. 1.8e-05;
  Matches   43;  Conservative   13;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              |:|| ||||||   ||| |: |||| || ||| | |::|:|| |||::| :||| ||
Db         52 GUGAGAGACGGAGGGCCCAUGGAACUUACAGAAGAAGUUGUUGAUGAGUUCAUGGAAGAU 111

Qy         61 GTACCAATGGCTGTGA 76
              |: || |:| |  : |
Db        112 GUCCCUAUGUCAAUCA 127


Search completed: December  2, 2004, 07:53:21
Job time : 410 secs
```

OM nucleic - nucleic search, using sw model

Run on: December 2, 2004, 05:41:26 ; Search time 2971 Seconds
(without alignments)
932.151 Million cell updates/sec

Title: US-09-551-494-5_COPY_5430_5505
Perfect score: 76
Sequence: 1 gtgacagacggctcgccaat.....tgaagtaccaatggctgtga 76

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	39.4	51.8	188	4	BM068137	BM068137 KS08017C0
c	2	39.4	51.8	348	4	BM067518	BM067518 KS08006E1
	3	32.6	42.9	583	6	CB264749	CB264749 41-E01466
	4	31.6	41.6	954	8	B12288	B12288 T2M2-Sp6 TA
	5	31	40.8	389	1	AV441961	AV441961 AV441961
	6	31	40.8	530	6	CA963974	CA963974 cATI008D1
	7	31	40.8	591	9	FR0020421	AL013304 F.rubripe
	8	31	40.8	656	5	BU636714	BU636714 010F08 In
	9	31	40.8	686	5	BU636451	BU636451 052H05 In
	10	31	40.8	1535	3	CNS0ADUO	BX813913 Arabidops
	11	31	40.8	1579	3	CNS0ACVY	BX815501 Arabidops
c	12	30.2	39.7	711	5	BX845438	BX845438 BX845438
c	13	30.2	39.7	800	5	BU915084	BU915084 AGENCOURT
c	14	30	39.5	446	9	FR0020459	AL013342 F.rubripe
c	15	29.8	39.2	561	5	BX553578	BX553578 BX553578
	16	29.6	38.9	357	9	CG742490	CG742490 ZMMBBc019
	17	29.6	38.9	854	7	CO800453	CO800453 AGENCOURT
c	18	29.4	38.7	732	1	AU305571	AU305571 AU305571
	19	28.8	37.9	387	5	BU816956	BU816956 UA11BPB12
	20	28.8	37.9	398	9	CNS00STV	AL088849 Arabidops
c	21	28.8	37.9	460	7	CK095490	CK095490 UA11BPB12
c	22	28.8	37.9	498	8	AZ912881	AZ912881 RPCI-24-1
	23	28.8	37.9	596	1	AJ769214	AJ769214 AJ769214
	24	28.8	37.9	605	1	AJ774675	AJ774675 AJ774675
	25	28.8	37.9	643	1	AJ767586	AJ767586 AJ767586
c	26	28.8	37.9	946	8	CC099545	CC099545 CSU-K34.1
c	27	28.6	37.6	514	8	BZ590961	BZ590961 3590_1_80
c	28	28.6	37.6	516	8	BZ590922	BZ590922 3590_1_80
c	29	28.6	37.6	522	8	BZ590931	BZ590931 3590_1_80
c	30	28.6	37.6	527	8	BZ588716	BZ588716 3590_1_9_
c	31	28.6	37.6	529	8	BZ586702	BZ586702 3590_1_20
	32	28.6	37.6	548	7	CN947691	CN947691 020802AVB
c	33	28.6	37.6	594	8	BZ586917	BZ586917 3590_1_22
c	34	28.6	37.6	763	9	CC621362	CC621362 OGUEX32TH
c	35	28.6	37.6	877	9	CG305892	CG305892 OGWMC12TV
c	36	28.6	37.6	1003	8	BZ700933	BZ700933 PUCEH52TD
	37	28.2	37.1	385	9	CG742799	CG742799 ZMMBBc019
	38	28.2	37.1	616	9	CC681582	CC681582 OGUCZ39TH
	39	28.2	37.1	695	9	CG392856	CG392856 ZMMBBc000
	40	28.2	37.1	938	9	CL238374	CL238374 ZMMBBb058
	41	28	36.8	244	9	CL226782	CL226782 ZMMBBc000
	42	28	36.8	264	9	CG838137	CG838137 ZMMBBc021
	43	28	36.8	276	9	CG743187	CG743187 ZMMBBc019
	44	28	36.8	300	9	CG831760	CG831760 ZMMBBc010
	45	28	36.8	547	6	CD836299	CD836299 BN45.048F

ALIGNMENTS

RESULT 1
BM068137/c
LOCUS BM068137 188 bp mRNA linear EST 11-SEP-2002
DEFINITION KS08017C07 KS08 Capsicum annuum cDNA, mRNA sequence.
ACCESSION BM068137
VERSION BM068137.1 GI:22788242
KEYWORDS EST.
SOURCE Capsicum annuum
ORGANISM Capsicum annuum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiids; Solanales; Solanaceae; Capsicum.
REFERENCE 1 (bases 1 to 188)
AUTHORS Lee,S., Kim,S.-Y., Chung,Y.-H., Shin,H.-J., Goh,S.-H., Pai,H.-S., Hur,C.-G. and Choi,D.
TITLE Generation of Expressed Sequence Tags from Hot Pepper (Capsicum annuum L.) and Sequence Analysis in Relation to Hypersensitive Response Against Pathogen
JOURNAL Unpublished (2001)
COMMENT Contact: Doil Choi
Genome Research Center and National Center for Genome Information
Korea Research Institute of Bioscience and Biotechnology
P.O. Box 115, Yusong, Taejeon, 305-600, Republic of Korea
Tel: 82-42-860-4340
Fax: 82-42-860-4309
Email: doil@mail.kribb.re.kr
High quality sequence stop: 188.
FEATURES Location/Qualifiers
source 1. .188
/organism="Capsicum annuum"
/mol_type="mRNA"
/cultivar="Hang Keun"
/db_xref="taxon:4072"
/tissue_type="anther"
/dev_stage="10 weeks after germination"
/clone_lib="KS08"
/note="Vector: pBluescript SK(-)"
ORIGIN

Query Match 51.8%; Score 39.4; DB 4; Length 188;
Best Local Similarity 71.2%; Pred. No. 0.01;
Matches 52; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
Db 122 GTGTCGGAAGGAGGACCCGTTGAACTTACAGAAGCAGTTGTTGATGAGTTCATCGAATCA 63

```
Qy          61 GTACCAATGGCTG 73
               || ||||||||||
Db          62 GTTCCAATGGCTG 50


RESULT 2
BM067518/c
LOCUS       BM067518                348 bp    mRNA    linear   EST 11-SEP-2002
DEFINITION  KS08006E10 KS08 Capsicum annuum cDNA, mRNA sequence.
ACCESSION   BM067518
VERSION     BM067518.1  GI:22787638
KEYWORDS    EST.
SOURCE      Capsicum annuum
  ORGANISM  Capsicum annuum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Capsicum.
REFERENCE   1  (bases 1 to 348)
  AUTHORS   Lee,S., Kim,S.-Y., Chung,Y.-H., Shin,H.-J., Goh,S.-H., Pai,H.-S.,
            Hur,C.-G. and Choi,D.
  TITLE     Generation of Expressed Sequence Tags from Hot Pepper (Capsicum
            annuum L.) and Sequence Analysis in Relation to Hypersensitive
            Response Against Pathogen
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Doil Choi
            Genome Research Center and National Center for Genome Information
            Korea Research Institute of Bioscience and Biotechnology
            P.O. Box 115, Yusong, Taejeon, 305-600, Republic of Korea
            Tel: 82-42-860-4340
            Fax: 82-42-860-4309
            Email: doil@mail.kribb.re.kr
            High quality sequence stop: 348.
FEATURES             Location/Qualifiers
     source          1. .348
                     /organism="Capsicum annuum"
                     /mol_type="mRNA"
                     /cultivar="Hang Keun"
                     /db_xref="taxon:4072"
                     /tissue_type="anther"
                     /dev_stage="10 weeks after germination"
                     /clone_lib="KS08"
                     /note="Vector: pBluescript SK(-)"
ORIGIN

  Query Match             51.8%;  Score 39.4;  DB 4;  Length 348;
  Best Local Similarity   71.2%;  Pred. No. 0.011;
  Matches   52;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              ||| | || ||    ||  ||||||| || |||   |||||||| ||||| | ||   |
Db        294 GTGTCGGAAGGAGGACCCGTTGAACTTACAGAAGCAGTTGTTGATGAGTTCATCGAATCA 235

Qy         61 GTACCAATGGCTG 73
              || ||||||||||
Db        234 GTTCCAATGGCTG 222


RESULT 3
CB264749
LOCUS       CB264749                583 bp    mRNA    linear   EST 06-NOV-2003
DEFINITION  41-E014660-035-002-B11-T7R MPIZ-ADIS-035 Arabidopsis thaliana cDNA
            clone MPIZp2000B112Q 5-PRIME, mRNA sequence.
ACCESSION   CB264749
VERSION     CB264749.1  GI:32889522
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 583)
  AUTHORS   Schmid,K.J., Soerensen,T.R., Stracke,R., Torjek,O., Altmann,T.,
            Mitchell-Olds,T. and Weisshaar,B.
  TITLE     Large-scale identification and analysis of genome-wide
            single-nucleotide polymorphisms for mapping in Arabidopsis thaliana
  JOURNAL   Genome Res. 13 (6), 1250-1257 (2003)
  MEDLINE   22683290
   PUBMED   12799357
COMMENT     Contact: Weisshaar B
            ADIS DNA core facility at MPIZ
            Max-Planck-Institute for Plant Breeding Research
            Carl-von-Linne Weg 10, 50829 Koeln, Germany
            Fax: 00492215062851
            Email: weisshaa@mpiz-koeln.mpg.de
            Insert Length: 583   Std Error: 0.00
            Plate: 2  row: B  column: 11
            Seq primer: T7R; CTAATACGACTCACTATAGGGA.
FEATURES             Location/Qualifiers
     source          1. .583
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /cultivar="Achkarren-2 (Ak-2)"
                     /db_xref="GABI:594666"
                     /db_xref="taxon:3702"
                     /clone="MPIZp2000B112Q"
                     /tissue_type="inflorescence"
                     /lab_host="E. coli TOP10"
                     /clone_lib="MPIZ-ADIS-035"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; cDNA
                     library from Arabidopsis thaliana, accession Achkarren-2;
                     inflorescences from flower buds to young siliques; library
                     was made at the Max-Planck-Institute for Plant Breeding
                     Research, Cologne, Germany; cloning sites SalI-NotI,
                     primer sites and orientation:
                     T7-SalI-CCACGCGTCCG-5prime-cDNA-polyA-CC-NotI-SP6; Note:
                     Sequencing granted in the context of the GABI Arabidopsis
                     VerbundI: Genetic Diversity, 'Establishment of
                     high-efficiency SNP-based mapping tools and development of
                     methods for genome-wide mutation detection' PI: Bernd
                     Weisshaar; Sequence submission managed by
                     RZPD/GABI-Primary database: http://gabi.rzpd.de This clone
                     is available from RZPD; contact RZPD (clone@rzpd.de) for
                     further information."
ORIGIN

  Query Match             42.9%;  Score 32.6;  DB 6;  Length 583;
  Best Local Similarity   66.2%;  Pred. No. 2.6;
  Matches   47;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy          6 AGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAAGTACC 65
              ||| |  || ||| ||||| |||  || || | | |||| | || |  || ||| | |
Db        381 AGATGCTTCTCCACTTGAAATCATGGATAAAGCTCTTGAGAAATTTGGAGACGAAATCGC 440

Qy         66 AATGGCTGTGA 76
              ||| ||| | |
Db        441 AATTGCTTTCA 451


RESULT 4
B12288
LOCUS       B12288                  954 bp    DNA     linear   GSS 14-MAY-1997
DEFINITION  T2M2-Sp6 TAMU Arabidopsis thaliana genomic clone T2M2, genomic
            survey sequence.
ACCESSION   B12288
VERSION     B12288.1  GI:2093409
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 954)
  AUTHORS   Feng,J., Dewar,K., Buehler,E., Kim,C., Li,Y., Shinn,P., Sun,H. and
            Ecker,J.
  TITLE     BAC End Sequences at ATGC
  JOURNAL   Unpublished (1997)
```

```
COMMENT     Other_GSSs: T2M2-T7
            Contact: Ecker J.
            Arabidopsis Thaliana Genome Center
            University of Pennsylvania
            Dept. of Biology, University of Pennsylvania, Philadelphia, PA
            19104
            Tel: 215-898-9384
            Fax: 215-898-8780
            Email: jecker@atgenome.bio.upenn.edu
            Seq primer: Sp6
            Class: BAC ends
            High quality sequence start: 93
            High quality sequence stop: 103.
FEATURES             Location/Qualifiers
     source          1. .954
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /ecotype="Columbia"
                     /db_xref="taxon:3702"
                     /clone="T2M2"
                     /sex="hermaphrodite"
                     /clone_lib="TAMU"
                     /note="Vector: BeloBACII; Site_1: HindIII; Site_2:
                     HindIII; Produced by Rod Wing"
ORIGIN

  Query Match             41.6%;  Score 31.6;  DB 8;  Length 954;
  Best Local Similarity   64.8%;  Pred. No. 6.3;
  Matches   46;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          6 AGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAAGTACC 65
              | |  |||| ||| ||||| |||  || || | | |||||   ||||  ||  || | |
Db         72 ANAATGCTCTCCACTTGAAATCATGGATAAAGCTCTTGAGAGATTCGGAGAACAATTCGC 131

Qy         66 AATGGCTGTGA 76
              ||| ||| | |
Db        132 AATTGCTTTTA 142


RESULT 5
AV441961
LOCUS       AV441961                389 bp    mRNA    linear   EST 18-FEB-2004
DEFINITION  AV441961 Arabidopsis thaliana above-ground organ two to six-week
            old Arabidopsis thaliana cDNA clone APD19f09_r 5', mRNA sequence.
ACCESSION   AV441961
VERSION     AV441961.1  GI:7612359
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 389)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     A large scale analysis of cDNA in Arabidopsis thaliana: Generation
            of 12,028 non-redundant expressed sequence tags from normalized and
            size-selected cDNA libraries
  JOURNAL   DNA Res. 7 (3), 175-180 (2000)
  MEDLINE   20363093
   PUBMED   10907847
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /ecotype="Columbia"
                     /db_xref="taxon:3702"
                     /clone="APD19f09_r"
                     /tissue_type="aboveground organs"
                     /dev_stage="two to six-week old"
                     /clone_lib="Arabidopsis thaliana above-ground organ two to
                     six-week old"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI"
ORIGIN

  Query Match             40.8%;  Score 31;  DB 1;  Length 389;
  Best Local Similarity   64.8%;  Pred. No. 8.5;
  Matches   46;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          6 AGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAAGTACC 65
              ||| |  || ||| ||||| |||  || || | | |||||   ||||  ||  || | |
Db         62 AGATGCTTCTCCACTTGAAATCATGGATAAAGCTCTTGAGAGATTCGGAGACCAAATCGC 121

Qy         66 AATGGCTGTGA 76
              ||| ||| | |
Db        122 AATTGCTTTTA 132


RESULT 6
CA963974
LOCUS       CA963974                530 bp    mRNA    linear   EST 03-JAN-2003
DEFINITION  cATI008D10AF Infected Arabidopsis Leaf Arabidopsis thaliana cDNA,
            mRNA sequence.
ACCESSION   CA963974
VERSION     CA963974.1  GI:27490531
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 530)
  AUTHORS   Lundsgaard,M., Emmersen,J., Nielsen,K.L., Wilson,I., Somerville,S.
            and Welinder,K.G.
  TITLE     EST sequencing of Erysiphe cichoracearum infected Arabidopsis
            plants
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Karen G. Welinder
            Institut for bioteknologi
            Aalborg Universitet
            Sohngaardsholmsvej 49, 9000 Aalborg, Denmark
            Tel: +45 96358467
            Fax: +45 98141808
            Email: kgw@bio.auc.dk.
FEATURES             Location/Qualifiers
     source          1. .530
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /ecotype="Columbia"
                     /db_xref="taxon:3702"
                     /dev_stage="Plant 3 weeks old, three days post infection"
                     /clone_lib="Infected Arabidopsis Leaf"
                     /note="Organ: Leaf; Vector: pBluescript; Mixed cDNA
                     library of Arabidopsis and E. cichoracearum infected leaf
                     from three weeks old Arabidopsis plants. Plants were
                     harvested 3 days after infection and mRNA oligo dT
                     selected."
ORIGIN

  Query Match             40.8%;  Score 31;  DB 6;  Length 530;
  Best Local Similarity   64.8%;  Pred. No. 9;
  Matches   46;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          6 AGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAAGTACC 65
              ||| |  || ||| ||||| |||  || || | | |||||   ||||  ||  || | |
Db        215 AGATGCTTCTCCACTTGAAATCATGGATAAAGCTCTTGAGAGATTCGGAGACCAAATCGC 274

Qy         66 AATGGCTGTGA 76
              ||| ||| | |
Db        275 AATTGCTTTTA 285
```

RESULT 7
FR0020421
LOCUS FR0020421 591 bp DNA linear GSS 25-FEB-2004
DEFINITION F.rubripes GSS sequence, clone 041P11dB11, genomic survey sequence.
ACCESSION AL013304
VERSION AL013304.1 GI:2679672
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetradontoidea; Tetraodontidae; Takifugu.
REFERENCE 1
AUTHORS Elgar,G., Clark,M.S., Meek,S., Smith,S., Warner,S., Edwards,Y.J., Bouchireb,N., Cottage,A., Yeo,G.S., Umrania,Y., Williams,G. and Brenner,S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE 2 (bases 1 to 591)
AUTHORS Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrania,Y., Williams,G. and Brenner,S.
TITLE Direct Submission
JOURNAL Submitted (08-DEC-1997) MRC Human Genome Mapping Project Resource Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmp.mrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid
PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic sequence.
FEATURES Location/Qualifiers
source 1. .591
/organism="Takifugu rubripes"
/mol_type="genomic DNA"
/db_xref="taxon:31033"
/clone="041P11dB11"
/clone_lib="cosmid 041P11"
ORIGIN

Query Match 40.8%; Score 31; DB 9; Length 591;
Best Local Similarity 64.2%; Pred. No. 9.2;
Matches 43; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

```
Qy         10 GGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAAGTACCAATG 69
              ||   || | ||||||||   ||| | || ||||| ||  | | ||| ||| |   || |
Db        113 GGGCTGCNAGTTGAACTCNTGGAANATGTNGTTGACGATCTTGGGGACGAACTTGTAAAG 172

Qy         70 GCTGTGA 76
              | ||| |
Db        173 GTTGTCA 179
```

RESULT 8
BU636714
LOCUS BU636714 656 bp mRNA linear EST 23-SEP-2002
DEFINITION 010F08 Infected Arabidopsis Leaf Arabidopsis thaliana cDNA, mRNA sequence.
ACCESSION BU636714
VERSION BU636714.1 GI:23303969
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 656)
AUTHORS Lundsgaard,M., Emmersen,J., Nielsen,K.L., Wilson,I., Somerville,S. and Welinder,K.G.
TITLE EST sequencing of Erysiphe cichoracearum infected Arabidopsis plants
JOURNAL Unpublished (2002)
COMMENT Contact: Karen G. Welinder
Institut for bioteknologi
Aalborg Universitet
Sohngaardsholmsvej 49, 9000 Aalborg, Denmark
Tel: +45 96358467
Fax: +45 98141808
Email: kgw@bio.auc.dk.
FEATURES Location/Qualifiers
source 1. .656
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/dev_stage="Plant 3 weeks old, three days post infection"
/clone_lib="Infected Arabidopsis Leaf"
/note="Organ: Leaf; Vector: pBluescript; Mixed cDNA library of Arabidopsis and E. cichoracearum infected leaf from three weeks old Arabidopsis plants. Plants were harvested 3 days after infection and mRNA oligo dT selected."
ORIGIN

Query Match 40.8%; Score 31; DB 5; Length 656;
Best Local Similarity 64.8%; Pred. No. 9.4;
Matches 46; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

```
Qy          6 AGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAAGTACC 65
              ||| |  || ||| ||||| |||  || || | | |||||   ||||  ||  || | |
Db        363 AGATGCTTCTCCACTTGAAATCATGGATAAAGCTCTTGAGAGATTCGGAGACCAAATCGC 422

Qy         66 AATGGCTGTGA 76
              ||| ||| | |
Db        423 AATTGCTTTTA 433
```

RESULT 9
BU636451
LOCUS BU636451 686 bp mRNA linear EST 23-SEP-2002
DEFINITION 052H05 Infected Arabidopsis Leaf Arabidopsis thaliana cDNA, mRNA sequence.
ACCESSION BU636451
VERSION BU636451.1 GI:23303706
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 686)
AUTHORS Lundsgaard,M., Emmersen,J., Nielsen,K.L., Wilson,I., Somerville,S. and Welinder,K.G.
TITLE EST sequencing of Erysiphe cichoracearum infected Arabidopsis plants
JOURNAL Unpublished (2002)
COMMENT Contact: Karen G. Welinder
Institut for bioteknologi
Aalborg Universitet
Sohngaardsholmsvej 49, 9000 Aalborg, Denmark
Tel: +45 96358467
Fax: +45 98141808
Email: kgw@bio.auc.dk.
FEATURES Location/Qualifiers
source 1. .686
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/dev_stage="Plant 3 weeks old, three days post infection"
/clone_lib="Infected Arabidopsis Leaf"
/note="Organ: Leaf; Vector: pBluescript; Mixed cDNA

library of Arabidopsis and E. cichoracearum infected leaf from three weeks old Arabidopsis plants. Plants were harvested 3 days after infection and mRNA oligo dT selected."

ORIGIN

Query Match 40.8%; Score 31; DB 5; Length 686;
Best Local Similarity 64.8%; Pred. No. 9.5;
Matches 46; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

```
Qy          6 AGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAAGTACC 65
              ||| |  || ||| ||||| |||  || || | | |||||   ||||  ||  || | |
Db        347 AGATGCTTCTCCACTTGAAATCATGGATAAAGCTCTTGAGAGATTCGGAGACCAAATCGC 406

Qy         66 AATGGCTGTGA 76
              ||| || | |
Db        407 AATTGCTTTTA 417
```

RESULT 10
CNS0ADUO
LOCUS CNS0ADUO 1535 bp mRNA linear HTC 06-FEB-2004
DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTFB47ZC07 of Flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress).
ACCESSION BX813913
VERSION BX813913.1 GI:42474399
KEYWORDS HTC; GSLT_cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1535)
AUTHORS Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1535)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length librairies construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES Location/Qualifiers
source 1. .1535
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTFB47ZC07"
/tissue_type="Flowers and buds"
/plasmid="pCMVSPORT_6"
gene complement(1. .1535)
/gene="At1g62180"

ORIGIN

Query Match 40.8%; Score 31; DB 3; Length 1535;
Best Local Similarity 64.8%; Pred. No. 11;
Matches 46; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

```
Qy          6 AGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAAGTACC 65
              ||| |  || ||| ||||| |||  || || | | |||||   ||||  ||  || | |
Db        305 AGATGCTTCTCCACTTGAAATCATGGATAAAGCTCTTGAGAGATTCGGAGACCAAATCGC 364

Qy         66 AATGGCTGTGA 76
              ||| ||| | |
Db        365 AATTGCTTTTA 375
```

RESULT 11
CNS0ACVY
LOCUS CNS0ACVY 1579 bp mRNA linear HTC 06-FEB-2004
DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTLS68ZB02 of Adult vegetative tissue of strain col-0 of Arabidopsis thaliana (thale cress).
ACCESSION BX815501
VERSION BX815501.1 GI:42473038
KEYWORDS HTC; GSLT_cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1579)
AUTHORS Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1579)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length librairies construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES Location/Qualifiers
source 1. .1579
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTLS68ZB02"
/tissue_type="Adult vegetative tissue"
/plasmid="pCMVSPORT_6"
gene complement(1. .1579)
/gene="At1g62180"

ORIGIN

Query Match 40.8%; Score 31; DB 3; Length 1579;
Best Local Similarity 64.8%; Pred. No. 11;
Matches 46; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

```
Qy          6 AGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAAGTACC 65
              ||| |  || ||| ||||  |||  || |||| | |||||   ||||  ||  || | |
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 1, 2004, 20:23:09 ; Search time 1635 Seconds
(without alignments)
2198.176 Million cell updates/sec

Title: US-09-551-494-5_COPY_5430_5505
Perfect score: 76
Sequence: 1 gtgacagacggctcgccaat.....tgaagtaccaatggctgtga 76

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	76	100.0	6355	6	AX040177	AX040177 Sequence
2	76	100.0	6355	6	AX795380	AX795380 Sequence
3	76	100.0	6355	14	TMGCG	M34077 Tobacco mil
4	71.2	93.7	1413	14	TMGMPCPA	M34236 Tobacco mil
5	69.6	91.6	6356	14	AB078435	AB078435 Tobacco m
6	58.6	77.1	411	6	AX040184	AX040184 Sequence
7	45.8	60.3	912	14	ORVTRANS	X55296 Odontogloss
8	45.8	60.3	5997	6	E03624	E03624 DNA encodin
9	45.8	60.3	6597	6	E04305	E04305 cDNA encodi
10	45.8	60.3	6609	14	ORU34586	U34586 Odontogloss
11	45.8	60.3	6611	14	S83257	S83257 126 kda pro
12	45.8	60.3	6612	14	AY571290	AY571290 Odontoglo
13	44	57.9	1004	14	TOTMV2	V01406 TMV 3' end.
14	44	57.9	1019	6	E00088	E00088 Genomic RNA
15	44	57.9	6395	14	D63809	D63809 Tabacco mos
16	42.4	55.8	396	14	TMO307579	AJ307579 Tobacco m
17	42.4	55.8	632	6	A68760	A68760 Sequence 1
18	42.4	55.8	706	14	TMO308689	AJ308689 Tobacco m
19	42.4	55.8	750	14	TMO509080	AJ509080 Tobacco m
20	42.4	55.8	750	14	TMO509084	AJ509084 Tobacco m
21	42.4	55.8	782	14	TMO308692	AJ308692 Tobacco m
22	42.4	55.8	788	14	TMO308685	AJ308685 Tobacco m
23	42.4	55.8	790	14	TMO307583	AJ307583 Tobacco m
24	42.4	55.8	804	6	CQ793025	CQ793025 Sequence
25	42.4	55.8	804	14	TMO307582	AJ307582 Tobacco m
26	42.4	55.8	806	14	TMO308682	AJ308682 Tobacco m
27	42.4	55.8	806	14	TMO308684	AJ308684 Tobacco m
28	42.4	55.8	807	6	BD263879	BD263879 Viral exp
29	42.4	55.8	807	6	BD263880	BD263880 Viral exp
30	42.4	55.8	807	6	AR435699	AR435699 Sequence
31	42.4	55.8	807	6	AR435700	AR435700 Sequence
32	42.4	55.8	807	6	AX045758	AX045758 Sequence
33	42.4	55.8	807	6	AX592974	AX592974 Sequence
34	42.4	55.8	807	14	TMO307581	AJ307581 Tobacco m
35	42.4	55.8	807	14	TMO308683	AJ308683 Tobacco m
36	42.4	55.8	807	14	TMO308693	AJ308693 Tobacco m
37	42.4	55.8	807	14	TMO509081	AJ509081 Tobacco m
38	42.4	55.8	808	14	TMO307578	AJ307578 Tobacco m
39	42.4	55.8	808	14	TMO308686	AJ308686 Tobacco m
40	42.4	55.8	808	14	TMO308688	AJ308688 Tobacco m
41	42.4	55.8	808	14	TMO308690	AJ308690 Tobacco m
42	42.4	55.8	809	14	TMO308691	AJ308691 Tobacco m
43	42.4	55.8	811	14	TMO509082	AJ509082 Tobacco m
44	42.4	55.8	815	14	AY300161	AY300161 Tobacco m
45	42.4	55.8	817	14	TMO509083	AJ509083 Tobacco m

ALIGNMENTS

RESULT 1
AX040177
LOCUS AX040177 6355 bp DNA linear PAT 18-NOV-2000
DEFINITION Sequence 5 from Patent WO0063397.
ACCESSION AX040177
VERSION AX040177.1 GI:11230127
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Meulewaeter,F., Cornelisse, Jacobs,J., van Eldik,G. and Metzlaff,M.
TITLE Methods and means for delivering inhibitory rna to plants and applications thereof
JOURNAL Patent: WO 0063397-A 5 26-OCT-2000;
Aventis CropScience N.V. (BE)
FEATURES Location/Qualifiers
source 1. .6355
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="cDNA copy of the nucleotide sequence of the genome of TMV-U2"
ORIGIN

Query Match 100.0%; Score 76; DB 6; Length 6355;
Best Local Similarity 100.0%; Pred. No. 2.2e-15;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5430 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 5489

Qy         61 GTACCAATGGCTGTGA 76
              ||||||||||||||||
Db       5490 GTACCAATGGCTGTGA 5505
```

RESULT 2
AX795380
LOCUS AX795380 6355 bp mRNA linear PAT 04-OCT-2003
DEFINITION Sequence 7 from Patent WO03052108.

```
ACCESSION   AX795380
VERSION     AX795380.1  GI:37516053
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Metzlaff,M.H., Gossele,V.M., Meulewaeter,F. and Fache,I.C.
  TITLE     Improved methods and means for delivering inhibitory rna to plants
            and applications thereof
  JOURNAL   Patent: WO 03052108-A 7 26-JUN-2003;
            Bayer BioScience N.V. (BE)
FEATURES             Location/Qualifiers
     source          1. .6355
                     /organism="synthetic construct"
                     /mol_type="mRNA"
                     /db_xref="taxon:32630"
                     /note="cDNA sequence of the genome of TMV-U2"
ORIGIN

  Query Match           100.0%;  Score 76;  DB 6;  Length 6355;
  Best Local Similarity  100.0%;  Pred. No. 2.2e-15;
  Matches   76;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5430 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 5489

Qy         61 GTACCAATGGCTGTGA 76
              ||||||||||||||||
Db       5490 GTACCAATGGCTGTGA 5505


RESULT 3
TMGCG
LOCUS       TMGCG                   6355 bp ss-RNA     linear   VRL 03-AUG-1993
DEFINITION  Tobacco mild green mosaic virus complete genome.
ACCESSION   M34077 M22483
VERSION     M34077.1  GI:335243
KEYWORDS    coat protein.
SOURCE      Tobacco mild green mosaic virus
  ORGANISM  Tobacco mild green mosaic virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE   1  (bases 6127 to 6355)
  AUTHORS   Garcia-Arenal,F.
  TITLE     Sequence and structure at the genome 3' end of the U2-strain of
            tobacco mosaic virus, a histidine-accepting tobamovirus
  JOURNAL   Virology 167 (1), 201-206 (1988)
  MEDLINE   89045644
   PUBMED   3188396
REFERENCE   2  (bases 1 to 6355)
  AUTHORS   Solis,I. and Garcia-Arenal,F.
  TITLE     The complete nucleotide sequence of the genomic RNA of the
            tobamovirus tobacco mild green mosaic virus
  JOURNAL   Virology 177 (2), 553-558 (1990)
  MEDLINE   90320127
   PUBMED   2371769
COMMENT     Original source text: Tobacco mild green mosaic virus (strain
            U2-TMV), cDNA to viral RNA, from N.tabacum cv. Samsum.
            Draft entry and computer-readable sequence for [1] kindly submitted
            by F.Garcia-Arenal, 10-FEB-1989.  The RNA appears to have a tRNA-
            like, L-shaped structure at the 3' terminus, linked to a quasi-
            continuous double-helical stalk, with five pseudoknots involved in
            the formation of the whole structure.  However, the structure of
            U2-TMV RNA is less stringently conserved than the 3' termini of
            'vulgare' and other histidine-accepting tobamoviruses. Draft entry
            and computer-readable sequence for [1] kindly submitted by
            F.Garcia-Arenal, 08-MAY-1990, for release after publication.
FEATURES             Location/Qualifiers
     source          1. .6355
                     /organism="Tobacco mild green mosaic virus"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:12241"
CDS                  join(71. .3403,3407. .4900)
                     /note="183 kDa protein"
                     /codon_start=1
                     /protein_id="AAA47934.1"
                     /db_xref="GI:335244"
                     /translation="MAHIQSIISNALLESVSGKNTLVNDLARRRMYDTAVEEFNARDR
                     RPKVNFSKTISEEQTLLVSNAYPEFQITFYNTQNAVHSLAGGLRALELEYLMLQVPYG
                     SPTYDIGGNFAAHLFKGRDYVHCCMPNLDIRDIMRHEGQKDSIEMYLSRLSRSNKVIP
                     EFQREAFNRYAEAPNEVCCSKTFQDCRIHPPENSGRRYAVALHSLYDIPVHEFGAALI
                     SKNIHVCYAASILAEALLLDQTEVTLNEIGATFKREGDDVSFFFADESTLNYSHKYKN
                     ILHYVVKSYFPASSRIVYFKEFLVTRVNTWFCKFTKVDTYILYKSVRQVGCDSDQFYE
                     AMEDAFAYKKTLAMFNTERAIFRDTASVNFWFPPKMKDMVIVPLFEGSITSKKMTRSEV
                     IVNRDFVYTVLNHIRTYQAKALTYQNVLSFVESIRSRVIINGVTARSEWDVDKAILQP
                     LSMTFFLQTKLAALQDDIVMGKFRCLDKTTSELIWDEVGKFFGNVFPTIKERLVSRKI
                     LDVSENALKIKIPDLYVTWKDRFVAEYTKSEELPHLDIKKDLEEAEQMYDALSELSIL
                     KGADNFDIAKFKDMCKALDVSPDVAARVIVAVAENRSGLTLTFDKPTEENVAKALKST
                     ASEAVVCLEPTSEEVNVNKFSIAEKGRLPVCAESHGLTNANLEHQELESLNDFHKACV
                     DSVITKQMASVVYTGSLKVQQMKNYVDSLAASLSATVSNLCKSLKDEVGYDSDSREKV
                     GVWDVTLKKWLLKPAAKGHSWGVVLDYKGKMPTALLSYEGDRMVTESDWRRVAVSSDT
                     MVYSDIAKLQNLRKTMRDGEPHEPTAKMVLVDGVPGCGKYKGDFERFDLDEDLILVPG
                     KQAAAMIRRRANSSGLIRATMDNVRTVDSLLMHPKPRSHKRLFIDEGLMLHTGCVNFL
                     VLISGCDIAYIYGDTQQIPFINRVQNFPYPKHFEKLQVDEVEMRRTTLRCPGDVNFFL
                     QSKYEGAVTTTSTVQRSVSSEMIGGKGVLNSVSKPLKGKIVTFTQADKFELEEKGYKN
                     VNTVHEIQGETFEDVSLVRLTATPLTLISKSSPHVLVALTRHTKSFKYYTVVLDPLVQ
                     IISDLSSLSSFLLEMYMVEAGSRQLQMDAVFKGHNLFVATPKSGDFPDLQFYYDVCLP
                     GNSTILNKYDAVTMRLRDNSLNVKDCVLDFSKSIPMPKEVKPCLEPVLRTAAEPPRAA
                     GLLENLVAMIKRNFNAPDLTGTIDIESTASVVVDKFFDSYFIKKEKYTKNIAGVMTKD
                     SMMRWLENRKEVLLDDLANYNFTDLPAIDQYKHMIKAQPKQKLDLSIQNEYPALQTIV
                     YHSKQINGILAGFSELTRLLLEAFDSKKFLFFTRKTPEQIQEFFSDLDSHVPMDVLEL
                     DISKYDKSQNEFHCAVEYEIWKRLGLNEFLAEVWKQGHRKTTLKDYIAGIKTCLWYQR
                     KSGDVTTFIGNTVIIAACLGSMLPMEKVIKGAFCGDDSVLYFPKGLDFPDIQSCANLM
                     WNFEAKLYRKRYGYFCGRYIIHHDKGAIVYYDPLKLISKLGAKHIKDYDHLEELRVSL
                     CDVACSLGNWCLGFPQLNAAIKEVHKTAIDGSFAFNCVNKFLCDKFLFRTLFLNGC"
CDS                  71. .3406
                     /note="126 kDa protein"
                     /codon_start=1
                     /protein_id="AAA47935.1"
                     /db_xref="GI:335245"
                     /translation="MAHIQSIISNALLESVSGKNTLVNDLARRRMYDTAVEEFNARDR
                     RPKVNFSKTISEEQTLLVSNAYPEFQITFYNTQNAVHSLAGGLRALELEYLMLQVPYG
                     SPTYDIGGNFAAHLFKGRDYVHCCMPNLDIRDIMRHEGQKDSIEMYLSRLSRSNKVIP
                     EFQREAFNRYAEAPNEVCCSKTFQDCRIHPPENSGRRYAVALHSLYDIPVHEFGAALI
                     SKNIHVCYAASILAEALLLDQTEVTLNEIGATFKREGDDVSFFFADESTLNYSHKYKN
                     ILHYVVKSYFPASSRIVYFKEFLVTRVNTWFCKFTKVDTYILYKSVRQVGCDSDQFYE
                     AMEDAFAYKKTLAMFNTERAIFRDTASVNFWFPPKMKDMVIVPLFEGSITSKKMTRSEV
                     IVNRDFVYTVLNHIRTYQAKALTYQNVLSFVESIRSRVIINGVTARSEWDVDKAILQP
                     LSMTFFLQTKLAALQDDIVMGKFRCLDKTTSELIWDEVGKFFGNVFPTIKERLVSRKI
                     LDVSENALKIKIPDLYVTWKDRFVAEYTKSEELPHLDIKKDLEEAEQMYDALSELSIL
                     KGADNFDIAKFKDMCKALDVSPDVAARVIVAVAENRSGLTLTFDKPTEENVAKALKST
                     ASEAVVCLEPTSEEVNVNKFSIAEKGRLPVCAESHGLTNANLEHQELESLNDFHKACV
                     DSVITKQMASVVYTGSLKVQQMKNYVDSLAASLSATVSNLCKSLKDEVGYDSDSREKV
                     GVWDVTLKKWLLKPAAKGHSWGVVLDYKGKMPTALLSYEGDRMVTESDWRRVAVSSDT
                     MVYSDIAKLQNLRKTMRDGEPHEPTAKMVLVDGVPGCGKYKGDFERFDLDEDLILVPG
                     KQAAAMIRRRANSSGLIRATMDNVRTVDSLLMHPKPRSHKRLFIDEGLMLHTGCVNFL
                     VLISGCDIAYIYGDTQQIPFINRVQNFPYPKHFEKLQVDEVEMRRTTLRCPGDVNFFL
                     QSKYEGAVTTTSTVQRSVSSEMIGGKGVLNSVSKPLKGKIVTFTQADKFELEEKGYKN
                     VNTVHEIQGETFEDVSLVRLTATPLTLISKSSPHVLVALTRHTKSFKYYTVVLDPLVQ
                     IISDLSSLSSFLLEMYMVEAGSR"
exon                 <71. .3403
                     /note="183 kDa protein"
exon                 3407. .>4900
                     /note="183 kDa protein"
CDS                  4890. .5660
                     /note="28.5 kDa protein"
                     /codon_start=1
                     /protein_id="AAA47936.1"
                     /db_xref="GI:335246"
                     /translation="MAVSLRDTVKISEFIDLSKQDEILPAFMTKVKSVRISTVDKIMA
                     VKNDSLSDVDLLKGVKLVKKGYVCLADLVVSGEWNLPDNCRGGVSVCIVDKRMKRSKE
                     ATLGAYHAPACKKNFSFKLIPNYSITSEDAEKHPWQVLVNIKGVAMEEGYCPLSLEFV
                     SICVVHKNNVRKGLRERILSVTDGSPIELTEKVVEEFVDEVPMAVKLEKVPENKKEMV
                     GNNVNNKKINNSGKKGFKIEEIEDNVSDDESIASSSTF"
CDS                  5666. .6145
```

```
                     /note="coat protein"
                     /codon_start=1
                     /protein_id="AAA47937.1"
                     /db_xref="GI:335247"
                     /translation="MPYTINSPSQFVYLSSAYADPVQLINLCTNALGNQFQTQQARTT
                     VQQQFADAWKPVPSMTVRFPASDFYVYRYNSTLDPLITALLNSFDTRNRIIEVDNQPA
                     PNTTEIVNATQRVDDATVAIRASINNLANELVRGTGMFNQAGFETASGLVWTTTPAT"
ORIGIN

  Query Match           100.0%;  Score 76;  DB 14;  Length 6355;
  Best Local Similarity  100.0%;  Pred. No. 2.2e-15;
  Matches   76;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5430 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 5489

Qy         61 GTACCAATGGCTGTGA 76
              ||||||||||||||||
Db       5490 GTACCAATGGCTGTGA 5505


RESULT 4
TMGMPCPA
LOCUS       TMGMPCPA                1413 bp ss-RNA     linear   VRL 03-AUG-1993
DEFINITION  Tobacco mild green mosaic virus movement and coat protein genes,
            complete cds.
ACCESSION   M34236
VERSION     M34236.1  GI:335248
KEYWORDS    coat protein; movement protein.
SOURCE      Tobacco mild green mosaic virus
  ORGANISM  Tobacco mild green mosaic virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE   1  (bases 1 to 1413)
  AUTHORS   Nejidat,A., Cellier,F., Holt,C.A., Gafny,R., Eggenberger,A.L. and
            Beachy,R.N.
  TITLE     Transfer of the movement protein gene between two tobamoviruses:
            influence on local lesion development
  JOURNAL   Virology 180 (1), 318-326 (1991)
  MEDLINE   91082424
   PUBMED   1984654
COMMENT     Original source text: Tobacco mild green mosaic virus (strain PV
            228), cDNA to viral RNA.
            Draft entry and computer-readable sequence for [1] kindly submitted
            by A.Nejidat, 11-MAY-1990.
FEATURES             Location/Qualifiers
     source          1. .1413
                     /organism="Tobacco mild green mosaic virus"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:12241"
     CDS             51. .851
                     /codon_start=1
                     /product="movement protein"
                     /protein_id="AAA47938.1"
                     /db_xref="GI:335249"
                     /translation="MAVSLRDTVKISEFINLSKQDEILPAFMTKVKSVRISTVDKIMA
                     VKNDSLSDVDLLKGVKLVKNGYVCLAGLVVSGEWNLPDNCRGGVSVCIVDKRMKRSNE
                     ATLGAYHAPACKKNFSFKLIPNYSITSEDAEKNPWQVLVNIKGVAMEEGYCPLSLEFV
                     SICVVHKNNVKKGLRERILRVTDDSPIELTEKVVEEFVDEVPMAVKLERFRKTKKGKK
                     RKKEKKKRVVGNSVNNKKINNSGKKGLKVEEIEDNVSDDESIASSSTF"
     CDS             857. .1336
                     /codon_start=1
                     /product="coat protein"
                     /protein_id="AAA47939.1"
                     /db_xref="GI:335250"
                     /translation="MPYTINSPSQFVYLSSAYADPVQLINLCTNALGNQFQTQQARTT
                     VQQQFADAWKPVPSMTVRFPASDFYVYRYNSTLDPLITALLNSFDTRNRIIEVDNQPA
                     PNTTEIVNATQRVDDATVAIRASINNLANELVRGTGMFNQAGFETASGLVWTTTPAT"
ORIGIN

  Query Match           93.7%;  Score 71.2;  DB 14;  Length 1413;
  Best Local Similarity  96.1%;  Pred. No. 7.6e-14;
  Matches   73;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              || ||||||| ||||||||||||||||||||||| ||||||||||||||||||||||||
Db        591 GTAACAGACGACTCGCCAATTGAACTCACTGAAAAAGTTGTTGAGGAGTTCGTGGATGAA 650

Qy         61 GTACCAATGGCTGTGA 76
              ||||||||||||||||
Db        651 GTACCAATGGCTGTGA 666


RESULT 5
AB078435
LOCUS       AB078435                6356 bp    RNA     linear   VRL 22-JAN-2002
DEFINITION  Tobacco mild green mosaic virus complete genome, strain:Japanese.
ACCESSION   AB078435
VERSION     AB078435.1  GI:18253266
KEYWORDS    .
SOURCE      Tobacco mild green mosaic virus
  ORGANISM  Tobacco mild green mosaic virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE   1
  AUTHORS   Okuno,T., Hamada,H., Takeuchi,S., Morishima,N., Yoshimoto,E. and
            Hikichi,Y.
  TITLE     Nucleotide sequence of the Japanese isolate of Tobacco mild green
            mosaic virus
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 6356)
  AUTHORS   Okuno,T., Hamada,H., Takeuchi,S., Morishima,N., Yoshimoto,E. and
            Hikichi,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JAN-2002) Tetsuro Okuno, Kyoto University;
            Kitashirakawa Oiwakecho, Sakyou 606-8502, Japan
            (E-mail:okuno@kais.kyoto-u.ac.jp, Tel:81-75-753-6131,
            Fax:81-75-753-6131)
FEATURES             Location/Qualifiers
     source          1. .6356
                     /organism="Tobacco mild green mosaic virus"
                     /virion
                     /mol_type="genomic RNA"
                     /strain="Japanese"
                     /db_xref="taxon:12241"
                     /note="RNA transcripts synthesized in vitro from the cDNA
                     clone of TMGMV-J direct systemic infection in Nicotiana
                     benthamiana-TMGMV-Japanese strain"
     CDS             72. .4901
                     /note="position 3405. .3407 is readthrough product"
                     /codon_start=1
                     /transl_except=(pos:3405. .3407,aa:OTHER)
                     /product="183kDa protein"
                     /protein_id="BAB83986.1"
                     /db_xref="GI:18253267"
                     /translation="MAHIQSTISNALLESVSGKNTLVNDLARRRMYDTAVDEFNARDR
                     RPKVNFPPKTISEEQTLLVTNAYPEFQITFYNTQNAVHSLAGGLRALELEYLMLQVPYG
                     SPTYDIGGNFAAHLFKGRDYVHCCMPNLDIRDIMRHEGQKDSIEMYLSRLSRSNKVIP
                     EFQRAAFNRYAEAPNEVCCSKTFQDCRIHPPENSGRRYAISLHSLYDIPVHEFGAALI
                     SKDIHICYAAFHFSEALLLDQTEVTLNEIGATFKREGDDVSFFFADESTLNYNHKYKN
                     ILHYVVKSYFPASSRIVYFKEFLVTRVNTWFCKFTKVDTYILYKSVRQVGCDSDQFYE
                     AMEDAFTYKKALAMFNTERAILRDTASVNFWFPKMKDMVIVPLPEGSITSKKMTRSEV
                     IVNRDFVYTVLNHIRTYQAKALTYQNVLSFVESIRSRVIINGVTARSEWDVDKAILQP
                     LSMTPFLQTKLAALQDDIVMGKFRCLDKTTSELIWDEVGKFFGNVFPTIKERLVSRKI
                     LDVSENALKIKIPDLYVTWKDRFVAEYTKSEELPHLDIKKDLEEAEQMYDALSELSIL
                     RDADNFDIAKFKDMCKALCVNPDVAARVIVAVAENRSGLTLTFDKPTEENVAKALKST
                     ASEAVVRLEPTSDEVMVNKFSIAERGKLPVCAESHGLTNAALEHQELESLNDFHKACV
                     DSVITKQMASVVYTGSLKVQQMKNYVDSLAASLSATVSNLCKSLKDVVGYDSDSREKV
                     GVWDVTLKKWLLKPAAKGHSWGVVLDYKGKMFTALLSYEGDRMVTESDWRRVAVSSDT
                     MVYSDIAKLQNLRKTMRDGEPHEPTAKMVLVDGVPGCGKTKEILERVDLDEDLILVPG
                     KQAAAMIRRRANSSGLIRATMDNVRTVDSFLMHPKPRTHKRLFIDEGLMLHTGCVNFL
                     VLISGCDIAYIYGDTQQIPFINRVQNFPYPKHFEKLQVDEVEMRRTTLRCPGDVNFFL
                     QSKYEGAVSTTSTVQRSVSSEMIGGKGVLNSVSKPLKGKIVTFTQADKFELEEKGYKN
                     VNTVHEIQGETFEDVSLVRLTATPLTLISKSSPHVLVALTRHTKSFKYYTVVLDPLVQ
                     IISDLSSLSSFLLEMYMVEAGGRXQLQMDAVFKGHNLFVATPKSGDFPDLQFYYDVCL
                     PGNSTILNKYDAVTMRLRDNSLNVKDCVLDFSKSIPMPKEVKPCLEPVLRTAAEPPRV
                     AGLLENLVAMIKRNFNAPDLTGTIDIESTASVVVDKFFDSYFLKKEKYTKNIAGVMTK
```

DSMMRWLGKQERSTIGQLANYNFVDLPAIDQYKHMIKAQPKQKLDLSIQNEYPALQTI
VYHSKQINGIFGPVFSELTRLLLEAVDSQKFLPFTRKTPEQIQEFFSDLDSHVPMDVL
ELDISKYDKSQNEFHCAVEYEIWKRLGLNEFLAEVWKQGHRKTTLKDYIAGIKTCLWY
QRKSGDVTTFIGNTVIIAACLGSMLPMEKVIKGAFCGDDSVLYFPKGLDFPDIQSCAN
LMWNFEAKLYRKKYGYFCGRYIIHHDKGAIVYYDPLKLISKLGAKHIKDYDHLEELRV
SLCDVACSLGNCAYFPQLNAAIKEVHKTAIDGSFAFNCVNKFLCDKFLFRTLFLNGC"

CDS 72. .3407
/codon_start=1
/product="126kDa protein"
/protein_id="BAB83987.1"
/db_xref="GI:18253268"
/translation="MAHIQSTISNALLESVSGKNTLVNDLARRRMYDTAVDEFNARDR
RPKVNFPKTISEEQTLLVTNAYPEFQITFYNTQNAVHSLAGGLRALELEYLMLQVPYG
SPTYDIGGNFAAHLFKGRDYVHCCMPNLDIRDIMRHEGQKDSIEMYLSRLSRSNKVIP
EFQRAAFNRYAEAPNEVCCSKTFQDCRIHPPENSGRRYAISLHSLYDIPVHEFGAALI
SKDIHICYAAFHFSEALLLDQTEVTLNEIGATFKREGDDVSFFFADESTLNYNHKYKN
ILHYVVKSYFPASSRIVYFKEFLVTRVNTWFCKFTKVDTYILYKSVRQVGCDSDQFYE
AMEDAFTYKKALAMFNTERAILRDTASVNFWFPKMKDMVIVPLFEGSITSKKMTRSEV
IVNRDFVYTVLNHIRTYQAKALTYQNVLSFVESIRSRVIINGVTARSEWDVDKAILQP
LSMTFFLQTKLAALQDDIVMGKFRCLDKTTSELIWDEVGKFFGNVFPTIKERLVSRKI
LDVSENALKIKIPDLYVTWKDRFVAEYTKSEELPHLDIKKDLEEAEQMYDALSELSIL
RDADNFDIAKFKDMCKALCVNPDVAARVIVAVAENRSGLTLTFDKPTEENVAKALKST
ASEAVVRLEPTSDEVMVNKFSIAERGKLPVCAESHGLTNANLEHQELESLNDFHKACV
DSVITKQMASVVYTGSLKVQQMKNYVDSLAAASLSATVSNLCKSLKDVVGYDSDSREKV
GVWDVTLKKWLLKPAAKGHSWGVVLDYKGKMPTALLSYEGDRMVTESDWRRVAVSSDT
MVYSDIAKLQNLRKTMRDGEPHEPTAKMVLVDGVPGCGKTKEILERVDLDEDLILVPG
KQAAAMIRRRANSSGLIRATMDNVRTVDSFLMHPKPRTHKRLFIDEGLMLHTGCVNFL
VLISGCDIAYIYGDTQQIPFINRVQNFPYPKHFEKLQVDEVEMRRTTLRCPGDVNFFL
QSKYEGAVSTTSTVQRSVSSEMIGGKGVLNSVSKPLKGKIVTFTQADKFELEEKGYKN
VNTVHEIQGETFEDVSLVRLTATPLTLISKSSPHVLVALTRHTKSFKYYTVVLDPLVQ
IISDLSSLSSFLLEMYMVEAGGR"

CDS 4891. .5661
/codon_start=1
/product="28kDa protein"
/protein_id="BAB83988.1"
/db_xref="GI:18253269"
/translation="MAVSLRDTVKISEFIDLSKQDEILPAFMTKVKSVRISTVDKIMA
VKNDSLSDVDLLKGVKLVKNGYVCLAGLVVSGEWNLPDNCRGGVSVCIVDKRMKRSKE
ATLGAYHAPACKKNFSFKLIPNYSITSEDAEKHPWQVLVNIKGVAMEEGYCPLSLEFV
SICVVHKNNVRKGLRERILRVTDGLPIELTEKVVEEFVDEVPMAVKLERFRKTKKRVV
GNSVNNKKINNSGKKGLKVEEIEDSVSDDESIASSSTF"

CDS 5667. .6146
/codon_start=1
/product="17kDa protein"
/protein_id="BAB83989.1"
/db_xref="GI:18253270"
/translation="MPYTINSPSQFVYLSSAYADPVQLINLCTNALGNQFQTQQARTT
VQQQFADAWKPVPSITVRFPASDFYVYRYNSTLDPLITALLNSFDTRNRIIEVDNQPA
PNTTEIVNATQRVDDATVAIRASINNLANELVRGTGMFNQAGFETASGLVWTTTPAT"

ORIGIN

Query Match 91.6%; Score 69.6; DB 14; Length 6356;
Best Local Similarity 94.7%; Pred. No. 3.7e-13;
Matches 72; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              || |||||||||| ||| |||||||||||||||| ||||||||||||||||||||||||
Db       5431 GTAACAGACGGCTTGCCGATTGAACTCACTGAAAAAGTTGTTGAGGAGTTCGTGGATGAA 5490

Qy         61 GTACCAATGGCTGTGA 76
              ||||||||||||||||
Db       5491 GTACCAATGGCTGTGA 5506
```

RESULT 6
AX040184
LOCUS AX040184 411 bp DNA linear PAT 18-NOV-2000
DEFINITION Sequence 12 from Patent WO0063397.
ACCESSION AX040184
VERSION AX040184.1 GI:11230134
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Meulewaeter,F., Cornelisse, Jacobs,J., van Eldik,G. and Metzlaff,M.
TITLE Methods and means for delivering inhibitory rna to plants and appl ications thereof
JOURNAL Patent: WO 0063397-A 12 26-OCT-2000;
Aventis CropScience N.V. (BE)
FEATURES Location/Qualifiers
source 1. .411
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="cDNA copy of part of the region of a TMV-U2 variant comprising the origin of assembly"
ORIGIN

Query Match 77.1%; Score 58.6; DB 6; Length 411;
Best Local Similarity 93.8%; Pred. No. 1.4e-09;
Matches 61; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy         12 CTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAAGTACCAATGGC 71
              |||||||||||||||||||||||| |||||||| ||||||| ||||||||||| |||||
Db          3 CTCGCCAATTGAACTCACTGAAAAAGTTGTTGATGAGTTCGTAGATGAAGTACCGATGGC 62

Qy         72 TGTGA 76
              |||||
Db         63 TGTGA 67
```

RESULT 7
ORVTRANS
LOCUS ORVTRANS 912 bp RNA linear VRL 31-DEC-1990
DEFINITION Odontoglossum ringspot virus cell-to-cell transport gene.
ACCESSION X55296
VERSION X55296.1 GI:60833
KEYWORDS transport protein.
SOURCE Odontoglossum ringspot virus
ORGANISM Odontoglossum ringspot virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE 1
AUTHORS Isomura,Y., Matumoto,Y., Murayama,A., Chatani,M., Inouye,N. and Ikegami,M.
TITLE Nucleotide sequence of cell-to-cell transport protein gene of odontoglossum ringspot virus
JOURNAL Nucleic Acids Res. 18 (24), 7448 (1990)
MEDLINE 91081343
PUBMED 2259639
REFERENCE 2 (bases 1 to 912)
AUTHORS Ikegami,M.
TITLE Direct Submission
JOURNAL Submitted (29-OCT-1990) Ikegami M., Nodai Research Institute, Tokyo University of Agriculture, 1-1-1 Sakuragaoka, Setagaya-ku, Tokyo 156, JAPAN
FEATURES Location/Qualifiers
source 1. .912
/organism="Odontoglossum ringspot virus"
/mol_type="genomic RNA"
/strain="Japan isolate"
/db_xref="taxon:12238"
/clone="pORE5-25"
CDS 1. .912
/codon_start=1
/product="cell-to-cell transport protein"
/protein_id="CAA39008.1"
/db_xref="GI:60834"
/db_xref="GOA:P22590"
/db_xref="Swiss-Prot:P22590"
/translation="MGRLRFVVLLSIFPIKTFSEPCSTMALVLRDSIKISEFINLSAS
EKLLPSALTAVKSVRISKVDKIISYENDTLSDIDLLKGVKLVENGYVCLAGLVVTGEW
NLPDNCKGGVSICLVDKRMKRANEATLGSYHTSACKKRFTFKIIPNYSVTTADALKGI
WQVMTNIRGVEMEKGPCPLSLEFVSICVVYLNNIKLGLREKILNVTEGGPTELTEAVV
DEFVEKVPMAARLKSFRSVNKKKPSNSSKFVNGKSRLNSRNKLNYENGDSDVGISVVD
DIVVGNGVSDIRIDDDCESFDAQSDSY"

```
     mat_peptide     1. .909
                     /product="cell-to-cell transport protein"
ORIGIN

  Query Match             60.3%;  Score 45.8;  DB 14;  Length 912;
  Best Local Similarity   76.7%;  Pred. No. 4.8e-05;
  Matches   56;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              || ||||| ||   ||| |  ||||| ||||||   |||||||| |||||||||||  ||
Db        613 GTAACAGAAGGAGGGCCCACCGAACTTACTGAAGCAGTTGTTGATGAGTTCGTGGAGAAA 672

Qy         61 GTACCAATGGCTG 73
              || || |||||||
Db        673 GTTCCTATGGCTG 685


RESULT 8
E03624
LOCUS       E03624                  5997 bp    RNA     linear   PAT 29-SEP-1997
DEFINITION  DNA encoding a part of complementary DNA to odontoglossum ringspot
            virus genomic RNA.
ACCESSION   E03624
VERSION     E03624.1  GI:2171839
KEYWORDS    JP 1992144685-A/1.
SOURCE      Odontoglossum ringspot virus
  ORGANISM  Odontoglossum ringspot virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE   1  (bases 1 to 5997)
  AUTHORS   Isomura,Y., Matsumoto,Y., Chatani,M. and Ikegami,M.
  TITLE     CDNA OF ORSV GENE
  JOURNAL   Patent: JP 1992144685-A 1 19-MAY-1992;
            NIPPON OIL CO LTD
COMMENT     OS   odontoglossum ringspot virus
            PN   JP 1992144685-A/1
            PD   19-MAY-1992
            PF   28-NOV-1989 JP 1989306626
            PI   ISOMURA YOSHIKATSU, MATSUMOTO YOSHITOMO, CHATANI MASAAKI, PI
            IKEGAMI MASATO
            PC   C12N15/40,C07K13/00,C07K15/04,C12N1/21,C12N5/10,C12P21/02, PC
             (C12P21/02,
            PC   C12R1:19);
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            CC   *source: clone=pORE5-25;
            FH   Key             Location/Qualifiers
            FH
            FT   misc_feature    1. .5997
            FT                   /note='a part of complementary DNA to  FT
                       Odontoglossum
            FT                   ringspot virus genomic RNA'
            FT                   /note='a part of complementary DNA to ORSV'.
FEATURES             Location/Qualifiers
     source          1. .5997
                     /organism="Odontoglossum ringspot virus"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:12238"
ORIGIN

  Query Match             60.3%;  Score 45.8;  DB 6;  Length 5997;
  Best Local Similarity   76.7%;  Pred. No. 7.1e-05;
  Matches   56;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              || ||||| ||   ||| |  ||||| ||||||   |||||||| |||||||||||  ||
Db       4805 GTAACAGAAGGAGGGCCCACCGAACTTACTGAAGCAGTTGTTGATGAGTTCGTGGAGAAA 4864

Qy         61 GTACCAATGGCTG 73
              || || |||||||
Db       4865 GTTCCTATGGCTG 4877


RESULT 9
E04305
LOCUS       E04305                  6597 bp    RNA     linear   PAT 29-SEP-1997
DEFINITION  cDNA encoding Odontoglossum ring spot virus (ORSV) genomic RNA.
ACCESSION   E04305
VERSION     E04305.1  GI:2172508
KEYWORDS    JP 1993030975-A/1.
SOURCE      Odontoglossum ringspot virus
  ORGANISM  Odontoglossum ringspot virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE   1  (bases 1 to 6597)
  AUTHORS   Isomura,Y., Matsumoto,Y., Chatani,M., Mizuta,Y. and Ikegami,M.
  TITLE     CDNA OF ORSV GENE
  JOURNAL   Patent: JP 1993030975-A 1 09-FEB-1993;
            NIPPON OIL CO LTD
COMMENT     OS   Odontoglossum ring spot virus
            PN   JP 1993030975-A/1
            PD   09-FEB-1993
            PF   26-JUL-1991 JP 1991276075
            PI   ISOMURA YOSHIKATSU, MATSUMOTO YOSHITOMO, CHATANI MASAAKI, PI
            MIZUTA YOSHINORI, IKEGAMI MASATO
            PC   C12N15/40,C07K15/04,C12N1/21,C12N15/11,C12N15/70,C12P21/02, PC
              C12Q1/68,
            PC   (C12N1/21,C12R1:19),(C12N15/70,C12R1:19),(C12P21/02,C12R1:19);
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            CC   *source: clone=pORCP-D83;
            FH   Key             Location/Qualifiers
            FH
            FT   misc_feature    1. .6597
            FT                   /note='cDNA to Odontoglossum ring spot virus
            FT                   genomic RNA'
            FT   CDS             5707. .6183
            FT                   /product='coat protein'
            FT   mat_peptide     5707. .6180
            FT                   /product='coat protein'.
FEATURES             Location/Qualifiers
     source          1. .6597
                     /organism="Odontoglossum ringspot virus"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:12238"
ORIGIN

  Query Match             60.3%;  Score 45.8;  DB 6;  Length 6597;
  Best Local Similarity   76.7%;  Pred. No. 7.3e-05;
  Matches   56;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              || ||||| ||   ||| |  ||||| ||||||   |||||||| |||||||||||  ||
Db       5405 GTAACAGAAGGAGGGCCCACCGAACTTACTGAAGCAGTTGTTGATGAGTTCGTGGAGAAA 5464

Qy         61 GTACCAATGGCTG 73
              || || |||||||
Db       5465 GTTCCTATGGCTG 5477


RESULT 10
ORU34586
LOCUS       ORU34586                6609 bp    RNA     linear   VRL 09-JUL-1996
DEFINITION  Odontoglossum ringspot virus, complete genome.
ACCESSION   U34586
VERSION     U34586.1  GI:1407591
KEYWORDS    .
SOURCE      Odontoglossum ringspot virus
  ORGANISM  Odontoglossum ringspot virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE   1  (bases 1 to 6609)
  AUTHORS   Chng,C.G., Wong,S.M., Mahtani,P.H., Loh,C.S., Goh,C.J., Kao,M.C.,
            Chung,M.C. and Watanabe,Y.
  TITLE     The complete sequence of a Singapore isolate of odontoglossum
```

```
          ringspot virus and comparison with other tobamoviruses
  JOURNAL Gene 171 (2), 155-161 (1996)
   MEDLINE 96257213
    PUBMED 8666266
REFERENCE 2  (bases 1 to 6609)
  AUTHORS Chng,C.G. and Wong,S.-M.
  TITLE   Direct Submission
  JOURNAL Submitted (22-AUG-1995) Sek-Man Wong, Botany, NUS, Kent Ridge,
          Singapore, 0511, Singapore
FEATURES             Location/Qualifiers
     source          1. .6609
                     /organism="Odontoglossum ringspot virus"
                     /mol_type="genomic RNA"
                     /strain="Singapore 1"
                     /db_xref="taxon:12238"
     CDS             join(63. .3398,3402. .4901)
                     /codon_start=1
                     /product="183 kDa readthrough protein"
                     /protein_id="AAC55012.1"
                     /db_xref="GI:1407592"
                     /translation="MAHFQQTMNNKVIEAGMGRNSLINDLAQRRVYDNAVEELNHRSR
                     RPRVNFSKVISQEQIIQATNAYPEFEITFYNTQLAVHSMAGGLRALELEYLMMQIPFG
                     SITYDIGGNFSAHLYKGRDYVHCCMPNLDIRDVARHINQQDTVSTYLARLERSKRGLP
                     VFQQSAFNKYMNDPDAVCCDKRFQDCSYSAGLPGKTYAVGLHSIYDIPADEFGAALLR
                     KDVHICYAAFHFSENLLLETTSAPLDEIGATFYKSGDRLSFFFQNESTLNYEHSYKNV
                     IKYVCKTFFPASNRFVYHKEFMCTRVNTWFCKFTKVDTYFLFRGVYTRGEDSEQFYTA
                     MDEAWEYKKTLAMLNSERTIFRDRAAVNFWFPKVKDMVIVPLFDGSVTSGKMKRSEVM
                     VNKDFVYTVLNHIRTYQDKALTYKNVLSFVESIRSRVIINGVTARSEWDVDKSVLQAL
                     SMTFLLQTKLAEAKDQVVLKKFQKFDDTVTNLFWKQISDAVGDLFPSIKERLISGGFV
                     KVAEQSLQIKTPDEYITPADKLVMEYQATEELQHLDISKPLERAEKYYNALSELSVLK
                     ECDEFDITQFKNLCEEKDIAPDVVAKVIVPIMKNELTLPFKNPTPEALSDALSPLPKD
                     LDMRFCLLKLSTCAPFPSVKTLDSGLLPKQSYGDERQFESQSVVSVSDFHLKSVESVK
                     MKSMSSAVYTGPLKVQQMKNYMDYLSASISATVSNLCKVLKDVYGADPESAEKSGVYD
                     VVKGKWLLKPKDKCHAWGVAELNNGEKVIVLLEWADGFPICGDWRRVAVSSDSPIYSD
                     MGKLQTLLSCLKDGEPVLRMPKVTLVDGVPGCGKTKEILETVNFDEDLILVPGKEACK
                     MIIKRANKSGHVRATRDNVRTVDSFLMHLKPKTYNKLFIDEGLMLHTGCVNFLVALSH
                     CREAMVFGDAEQIPFINRVANFPYPKHFRYTCLYHREVRRLSLRCPADVTHFMNSKYD
                     GKVLCTNDVIRSVDAEVVRGKGVFNPKSKPLKGKIITFTQSDKAELKERGYEEVSTFG
                     EINTVHEIQGETFEDVSVVRLTPTPLELISKSSPHVLVALTRHTKSFKYYSVVLDPLV
                     KVWSDLSKVSDFILDMYKVDAGILQLQVGSIFKGENLFVPCPKSGYISDMQTYYDTLV
                     PGNSTILNEYDAVTMNLRENNLNVKDCTIDFSKSVSVPRQQQEFFTPAHRTAAERPRS
                     AGLLENLVAMIKRNFNSPDLTGILDIEDTAELVVNKFWDAYIIDELSGGNVTPMTSDA
                     FHRWMAKQEKSTIGQLADFDFVDLPAIDQYKHMIKAQPKQKLGLSPQDEYAALQTIVY
                     HSKQINAIFGPLFSELTRQLLERIDSSKFLFYTRKTPEQIEAFFSDLDSTVPMEVLEL
                     DISKYDKSQNEFHCAVEYLIWEKLGLNGFLEEVWKQGHRKTSLKDYTAGIKTCLWYQR
                     KSGDVTTFIGNTVIIAACLASMIPMDKVIKAAFCGDDSILYIPKGLDLPDIQSGANLM
                     WNFEAKLYRKRYGYFCGRYIIHHDRGAIVYYDPLKLISKLGCKHIKSLDHLEEFRISL
                     CDVSSSLNNCAYFGQLNDAIAEVHKTAVNGSFAFCSIVKYLSDKNLFRTLFNNGSSTK
                     G"
     CDS             63. .3401
                     /codon_start=1
                     /product="126 kDa replicase"
                     /protein_id="AAC55013.1"
                     /db_xref="GI:1407593"
                     /translation="MAHFQQTMNNKVIEAGMGRNSLINDLAQRRVYDNAVEELNHRSR
                     RPRVNFSKVISQEQIIQATNAYPEFEITFYNTQLAVHSMAGGLRALELEYLMMQIPFG
                     SITYDIGGNFSAHLYKGRDYVHCCMPNLDIRDVARHINQQDTVSTYLARLERSKRGLP
                     VFQQSAFNKYMNDPDAVCCDKRFQDCSYSAGLPGKTYAVGLHSIYDIPADEFGAALLR
                     KDVHICYAAFHFSENLLLETTSAPLDEIGATFYKSGDRLSFFFQNESTLNYEHSYKNV
                     IKYVCKTFFPASNRFVYHKEFMCTRVNTWFCKFTKVDTYFLFRGVYTRGEDSEQFYTA
                     MDEAWEYKKTLAMLNSERTIFRDRAAVNFWFPKVKDMVIVPLFDGSVTSGKMKRSEVM
                     VNKDFVYTVLNHIRTYQDKALTYKNVLSFVESIRSRVIINGVTARSEWDVDKSVLQAL
                     SMTFLLQTKLAEAKDQVVLKKFQKFDDTVTNLFWKQISDAVGDLFPSIKERLISGGFV
                     KVAEQSLQIKTPDEYITPADKLVMEYQATEELQHLDISKPLERAEKYYNALSELSVLK
                     ECDEFDITQFKNLCEEKDIAPDVVAKVIVPIMKNELTLPFKNPTPEALSDALSPLPKD
                     LDMRFCLLKLSTCAPFPSVKTLDSGLLPKQSYGDERQFESQSVVSVSDFHLKSVESVX
                     MKSMSSAVYTGPLKVQQMKNYMDYLSASISATVSNLCKVLKDVYGADPESAEKSGVYD
                     VVKGKWLLKPKDKCHAWGVAELNNGEKVIVLLEWADGFPICGDWRRVAVSSDSPIYSD
                     MGKLQTLLSCLKDGEPVLRMPKVTLVDGVPGCGKTKEILETVNFDEDLILVPGKEACX
                     MIIKRANKSGHVRATRDNVRTVDSFLMHLKPKTYNKLFIDEGLMLHTGCVNFLVALSH
                     CREAMVFGDAEQIPFINRVANFPYPKHFRYTCLYHREVRRLSLRCPADVTHFMNSKYD
                     GKVLCTNDVIRSVDAEVVRGKGVFNPKSKPLKGKIITFTQSDKAELKERGYEEVSTFG
                     EINTVHEIQGETFEDVSVVRLTPTPLELISKSSPHVLVALTRHTKSFKYYSVVLDPLV
                     KVWSDLSKVSDFILDMYKVDAGIL"
     CDS             4807. .5718
                     /codon_start=1
                     /product="movement protein"
                     /protein_id="AAC55014.1"
                     /db_xref="GI:1407594"
                     /translation="MGRLRFVVLLSIFPIKTFSEPCSTMALVLRDSIKISEFINLSAS
                     EKLLPSALTAVKSVRISKVDKIISYENDTLSDIDLLKGVKLVENGYVCLAGLVVTGEW
                     NLPDNCKGGVSICLVDKRMKRANEATLGSYHTSACKKRFTFKIIPNYSVTTADALKGI
                     WQVMINIRGVEMEKGFCPLSLEFVSICVVYLNNIKLGLREKILNVTEGGPTELTEAVV
                     DEFVEKVPMAARLKSFRSVNKKKPSDSSKFVNGKSRLNSRNKLNYENGDSDVGISVVD
                     DIVVGNGVSDIRIDDDCESFDAQSDPY"
     CDS             5721. .6197
                     /codon_start=1
                     /product="coat protein"
                     /protein_id="AAC55015.1"
                     /db_xref="GI:1407595"
                     /translation="MSYTITDPSKLAYLSSAWADPNSLINLCTNSLGNQFQTQQARTT
                     VQQQFADVWQPVPTLTSRFPAGAGYFRVYRYDPILDPLITFLMGTFDTRNRIIEVENP
                     QNPTTTETLDATRRVDDATVAIRSAINNLLNELVRGTGMYNQVSFETMSGLTWTSS"
ORIGIN

  Query Match             60.3%;  Score 45.8;  DB 14;  Length 6609;
  Best Local Similarity   76.7%;  Pred. No. 7.3e-05;
  Matches   56;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              || ||||| ||   ||| |  ||||| ||||||   |||||||| ||||||||||  ||
Db       5419 GTAACAGAAGGAGGGCCCACCGAACTTACTGAAGCAGTTGTTGATGAGTTCGTGGAGAAA 5478

Qy         61 GTACCAATGGCTG 73
              || || |||||||
Db       5479 GTTCCTATGGCTG 5491


RESULT 11
S83257
LOCUS       S83257                  6611 bp    RNA     linear   VRL 12-MAR-1997
DEFINITION  126 kda protein. . .18 kda protein [Odontoglossum ringspot virus,
            Cy-1, Genomic RNA, 4 genes, 6611 nt].
ACCESSION   S83257
VERSION     S83257.1  GI:1881785
KEYWORDS    .
SOURCE      Odontoglossum ringspot virus
  ORGANISM  Odontoglossum ringspot virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE   1  (bases 1 to 6611)
  AUTHORS   Ikegami,M., Isomura,Y., Matumoto,Y., Chatani,M. and Inouye,N.
  TITLE     The complete nucleotide sequence of odontoglossum ringspot virus
            (Cy-1 strain) genomic RNA
  JOURNAL   Microbiol. Immunol. 39 (12), 995-1001 (1995)
  MEDLINE   96381046
   PUBMED   8789059
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 179070] from the original journal article.
FEATURES             Location/Qualifiers
     source          1. .6611
                     /organism="Odontoglossum ringspot virus"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:12238"
     gene            66. .4901
                     /gene="183 kda protein"
     CDS             join(66. .3398,3402. .4901)
                     /gene="183 kda protein"
                     /codon_start=1
                     /protein_id="AAB49498.1"
                     /db_xref="GI:1881786"
                     /translation="AHFQQTMNNKVIEAGMGRNSLINDLAQRRVYDNAVEELNHRSRR
                     PKVNFSKVISQEQIIQATNAYPEFEITFYNTQLAVHSMAGGLRALELEYLMMQIPFGS
                     ITYDIGGNFSAHLYKGRDYVHCCMPNLDIRDVARHINQQDTVSTYLARLERSKRGLPV
                     FQQSAFNKYMNDPDAVCCDKRFQDCSYSVDLPGKTYAVALHSIYDIPADEFGAALLRK
                     DVHICYAAFHFSENLLLETTSAPLDEIGATFYKSGDRLSFFFQNESTLNYEHSYKNVI
                     KYVCKTFFPASNRFVYHKEFMCTRVNTWFCKFTKVDTYFLFRGVYTRGEDSEQFYTAM
```

```
                     DEAWEYKKTLAMLKCERTIFRDRAAVNFWFPKVKDMVIVPLFDGSVTSGKMKRSEVMV
                     NKDFVYTVLNHIRTYQDKALTYKNVLSFVESIRSRVIINGVSARSEWDVDKSVLQALS
                     MTFLLQTKLAEAKDQVVLKKFQKFDDTVTNLFWKQISDAVGDLFPSIKETLISGGFVK
                     VAEQSLQIKTPDEYITFADKLVMEYKATEELQHLDISKPLERAEKYYNALSELSVLKE
                     CDEFDITQFKNLCEEKDIDPDVVAKVIVAIMKNELTLPFKNPTPEALSDALSPLPKDL
                     DMRFDLLKLSTCAPFPSVKTLDSGLLPKQSYGDERQFESQSVVSVSDFHLKSVESVKM
                     KSMSSAVYTGPLKVQQMKNYMDYLSASISATVSNLCKVLKDVYGADPESAEKSGVYDV
                     VKGKWLLKPKGKCHAWGVAELNNGEKVIVLLEWADGFPICGDWRRVAVSSDSLIYSDM
                     GKLQTLLSCLKDGEPVPSDAKVTLVDGVPGCGKTKEILETVNFDEDLILVPGKEACKM
                     IIKRANKSGHVRATKDNVRTVDSFLMHLKPKTYNKLFIDEGLMLHTGCVNFLIALSHC
                     REAMVFGDTEQIPFINRVANFPYPKHFATLVYDHREVRRLSLRCPADVTHFMNSKYDG
                     KVLCTNDVIRSVDAEVVRGKGVFNPKSKPLKGKIITFTQSDKAELKERGYEEVSTFGE
                     INTVHEIQGETFEDVSVVRLTPTPLELISKSSPHVLVALTRHTKSFKYYSVVLDPLVK
                     VCSDLSKVSDFILDMYKVDAGILQLQVGSIFKGENLFVPCPKSGYISDMQFYYDTLLP
                     GNSTILNEYDAVTMNLRENNLNVKDCTIDFSKSVSVPRQQQEFFTPVIRTAAERPRSR
                     GLLENLVAMIKRNFNSPDLTGILDIEDTAELVVNKFWDAYIIDELSGGNVTPMTSDAF
                     HRWMAKQEKSTIGQLADFDFVDLPAIDQYKHMIKAQPKQKLDLSPQDEYAALQTIVYH
                     SKQINAIFGPLFSELTRQLLERIDSSKFLFYTRKTPEQIEEFFSDLDSTVPMEVLELD
                     ISKYDKSQNEFHCAVEYLIWEKLGLNGFLEEVWKQGHRKTSLKDYTAGIKTCLWYQRK
                     SGDVTTFIGNTVIIAACLASMIPMDKVIKAAPCGDDSMLYIPKGLDLPDIQSGANLMW
                     NFEAKLYRKRYGYFCGRYIIHHDRGAIVYYDPVKLISKLGCKHIKSLDHLEEFRISLC
                     DVSASLNNCAYYGQLNDAIAEVHKTAVNGSFAFCSIVKYLSDKNLFRTLFYNGSSTKG
                     "
gene                 66. .3401
                     /gene="126 kda protein"
CDS                  66. .3401
                     /gene="126 kda protein"
                     /codon_start=1
                     /protein_id="AAB49499.1"
                     /db_xref="GI:1881787"
                     /translation="AHFQQTMNNKVIEAGMGRNSLINDLAQRRVYDNAVEELNHRSRR
                     PKVNFSKVISQEQIIQATNAYPEFEITFYNTQLAVHSMAGGLRALELEYLMMQIPFGS
                     ITYDIGGNFSAHLYKGRDYVHCCMPNLDIRDVARHINQQDTVSTYLARLERSKRGLPV
                     FQQSAFNKYMNDPDAVCCDKRFQDCSYSVDLPGKTYAVALHSIYDIPADEFGAALLRK
                     DVHICYAAFHFSENLLLETTSAPLDEIGATFYKSGDRLSFFFQNESTLNYEHSYKNVI
                     KYVCKTFFPASNRFVYHKEFMCTRVNTWFCKFTKVDTYFLFRGVYTRGEDSEQFYTAM
                     DEAWEYKKTLAMLKCERTIFRDRAAVNFWFPKVKDMVIVPLFDGSVTSGKMKRSEVMV
                     NKDFVYTVLNHIRTYQDKALTYKNVLSFVESIRSRVIINGVSARSEWDVDKSVLQALS
                     MTFLLQTKLAEAKDQVVLKKFQKFDDTVTNLFWKQISDAVGDLFPSIKETLISGGFVK
                     VAEQSLQIKTPDEYITFADKLVMEYKATEELQHLDISKPLERAEKYYNALSELSVLKE
                     CDEFDITQFKNLCEEKDIDPDVVAKVIVAIMKNELTLPFKNPTPEALSDALSPLPKDL
                     DMRFDLLKLSTCAPFPSVKTLDSGLLPKQSYGDERQFESQSVVSVSDFHLKSVESVKM
                     KSMSSAVYTGPLKVQQMKNYMDYLSASISATVSNLCKVLKDVYGADPESAEKSGVYDV
                     VKGKWLLKPKGKCHAWGVAELNNGEKVIVLLEWADGFPICGDWRRVAVSSDSLIYSDM
                     GKLQTLLSCLKDGEPVPSDAKVTLVDGVPGCGKTKEILETVNFDEDLILVPGKEACKM
                     IIKRANKSGHVRATKDNVRTVDSFLMHLKPKTYNKLFIDEGLMLHTGCVNFLIALSHC
                     REAMVFGDTEQIPFINRVANFPYPKHFATLVYDHREVRRLSLRCPADVTHFMNSKYDG
                     KVLCTNDVIRSVDAEVVRGKGVFNPKSKPLKGKIITFTQSDKAELKERGYEEVSTFGE
                     INTVHEIQGETFEDVSVVRLTPTPLELISKSSPHVLVALTRHTKSFKYYSVVLDPLVK
                     VCSDLSKVSDFILDMYKVDAGIL"
CDS                  4882. .5718
                     /gene="31 kda protein"
                     /codon_start=1
                     /protein_id="AAB49500.1"
                     /db_xref="GI:1881788"
                     /translation="ALVLRDSIKISEFINLSASEKLLPSALTAVKSVRISKVDKIISY
                     ENDTLSDIDLLKGVKLVENGYVCLAGLVVTGEWNLPDNCKGGVSICLVDKRMKRANEA
                     TLGSYHTSACKKRFTFKIIPNYSVTTADALKGIWQVMTNIRGVEMEKGFCPLSLEFVS
                     ICVVYLNNIKLGLREKILNVTEGGPTELTEAVVDEFVEKVPMAARLKSFRSVNKKKPS
                     NSSKFVNGKSRLNSRNKLNYENGDSDVGISVVDDIVVGNGVSDIRIDDDCESFDAQSD
                     SY"
gene                 4882. .5715
                     /gene="31 kda protein"
gene                 5724. .6197
                     /gene="18 kda protein"
CDS                  5724. .6197
                     /gene="18 kda protein"
                     /codon_start=1
                     /protein_id="AAB49501.1"
                     /db_xref="GI:1881789"
                     /translation="SYTITDPSKLAYLSSAWADPNSLINLCTNSLGNQFQTQQARTTV
                     QQQFADVWQPVPTLTSRFPAGAGYFRVYRYDPILDPLITFLMGTFDTRNRIIEVENPQ
                     NPTTTETLDATRRVDDATVAIRSAINNLLNELVRGTGMYNQVSFETISGLTWTSS"
```

```
ORIGIN

  Query Match           60.3%;  Score 45.8;  DB 14;  Length 6611;
  Best Local Similarity 76.7%;  Pred. No. 7.3e-05;
  Matches   56;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              || ||||| ||   ||| |  ||||| ||||||   |||||||| ||||||||||  ||
Db       5419 GTAACAGAAGGAGGGCCCACCGAACTTACTGAAGCAGTTGTTGATGAGTTCGTGGAGAAA 5478

Qy         61 GTACCAATGGCTG 73
              || || |||||||
Db       5479 GTTCCTATGGCTG 5491


RESULT 12
AY571290
LOCUS       AY571290                6612 bp    RNA     linear   VRL 14-APR-2004
DEFINITION  Odontoglossum ringspot virus strain Taiwan, complete genome.
ACCESSION   AY571290
VERSION     AY571290.1  GI:46309859
KEYWORDS    .
SOURCE      Odontoglossum ringspot virus
  ORGANISM  Odontoglossum ringspot virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE   1  (bases 1 to 6612)
  AUTHORS   Wang,H.L. and Wang,J.N.
  TITLE     Molecular sequencing and analysis of the viral genome of
            Odontoglossum ringspot virus Taiwan strain
  JOURNAL   Zhi Wu Bing Li Xue Hui Kan 13 (2004) In press
REFERENCE   2  (bases 1 to 6612)
  AUTHORS   Wang,H.L. and Wang,J.N.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAR-2004) Graduate Institute of Biology Science,
            National Kaohsiung Normal University, 116, Ho Ping 1 Rd, Kaohsiung
            802, Taiwan
FEATURES             Location/Qualifiers
     source          1. .6612
                     /organism="Odontoglossum ringspot virus"
                     /mol_type="genomic RNA"
                     /strain="Taiwan"
                     /db_xref="taxon:12238"
     gene            64. .4902
                     /gene="RdRp"
     CDS             64. .4902
                     /gene="RdRp"
                     /codon_start=1
                     /transl_except=(pos:3400. .3402,aa:OTHER)
                     /product="RNA-dependent RNA polymerase readthrough
                     protein"
                     /protein_id="AAS87224.1"
                     /db_xref="GI:46309861"
                     /translation="MAHFQQTMNNKVIEAGMGRNSLINDLAQRRVYDNAVEELNHRSR
                     RPKVNFSKVISQEQIIQATNAYPEFEITFYNTQLAVHSMAGGLRALELEYLMMQIPFG
                     SVTYDIGGNFSAHLYKGRDYVHCCMPNLDIRDVARHINQQDTVSTYLARLERSKRGLP
                     VPQQSAFNKYMNDPDAVCCDKRFQDCSYSVDLPGKTYAVALHSIYDIPADEFGAALLR
                     KDVHICYAAFHFSENLLLETTSAPLDEIGATFYKSGDRLSFFFQNESTLNYEHSYKNV
                     IKYVCKTFFPASNRFVYHKEFMCTRVNTWFCKFTKVDTYFLFRGVYTRGEDSEQFYTA
                     MDEAWEYKKTLAMLNSERTIFRDRAAVNFWFPKVKDMVIVPLFDGSVTSGKMKRSEVM
                     VNKDFVYTVLNHIRTYQDKALTYKNVLSFVESIRSRVIINGVTARSEWDVDKSVLQAL
                     SMTFLLQTKLAEAKDQVVLKKFQKFDDTVTNLFWKQISDAVGDLFPSIKERLISGGFV
                     KVAEQSLQIKTPDEYITFADKLVMEYKATEELQHLDISKPLERAEKYYNALSELSVLK
                     ECDEFDITQFKNLCEEKDIDPDVVAKVIVAIMKNELTLPFKNPTPEALSDALSPLPKD
                     LDMRFDLLKLSTCAPFPSVKTLDSGLLPKQSYGDERQFESQSVVSVSDFHLKSVESVK
                     MKNMSSAVYTGPLKVQQMKNYMDYLSASISATVSNLCKVLKDVYGADPESAEKSGVYD
                     VVKGKWLLKPKDKCHAWGVAELNNGEKVIVLLEWADGFPICGDWRRVAVSSDSLIYSD
                     MGKLQTLLSCLKDGEPVPSDAKVTLVDGVPGCGKTKEILETVNFDEDLILVPGKEACK
                     MIIKRANKSGHVRATKDNVRTVDSFLMHLKPKTYNKLFIDEGLMLHTGCVNFLIALSH
                     CREAMVFGDTEQIPFINRVANFPYPKHFATLVYDHREVRRLSLRCPADVTHFMNSKYD
                     GKVLCTSDVIRSVDAEVVRGKGVFNPKSKPLKGKIITFTQSDKAELKERGYEEVSTFG
                     EVNTVHEIQGETFEDVSVVRLTPTPLELISKSSPHVLVALTRHTKSFKYYSVVLDPLV
                     KVCSDLSKVSDFILDMYKVDAGILXQLQVGSIFKGENLFVPCPKSGYISDMQFYYDTL
                     LPGNSTILNEYDAVTMNLRENNLNVKDCTIDFSKSVSVPRQQQEFFTPVIRTAAERPR
```

```
                     SAGLLENLVAMIKRNFNSPDLTGILDIEDTAELVVNKFWDAYIIDELSGGNVTPMTSD
                     AFHRWMAKQEKSTIGQLADFDFVDLPAIDQYKHMIKAQPKQKLDLSPQDEYAALQTIV
                     YHSKQINAIFGPLFSELTRQLLERIDSSKPLFYTRKTPEQIEEFFSDLDSTVPMEVLE
                     LDISKYDKSQNEPHCAVEYLIWEKLGLNGFLEEVWKQGHRKTSLKDYTAGIKTCLWYQ
                     RKSGDVTTFIGNTVIIAACLASMIPMDKVIKAAFCGDDSILYIPKGLDLPDIQSGANL
                     MWNFEAKLYRKRYGYFCGRYIIHHDRGAIVYYDPLKLISKLGCKHIKSLDHLEEFRIS
                     LCDVSSSLNNCAYFGQLNDAIAEVHKTAVNGSFAFCSIVKYLSDKNLFRTLFYNGSST
                     KG"
     CDS             64. .3402
                     /gene="RdRp"
                     /codon_start=1
                     /product="RNA-dependent RNA polymerase/replicase"
                     /protein_id="AAS87223.1"
                     /db_xref="GI:46309860"
                     /translation="MAHFQQTMNNKVIEAGMGRNSLINDLAQRRVYDNAVEELNHRSR
                     RPKVNFSKVISQEQIIQATNAYPEFEITFYNTQLAVHSMAGGLRALELEYLMMQIPFG
                     SVTYDIGGNFSAHLYKGRDYVHCCMPNLDIRDVARHINQQDTVSTYLARLERSKRGLP
                     VFQQSAFNKYMNDPDAVCCDKRFQDCSYSVDLPGKTYAVALHSIYDIPADEFGAALLR
                     KDVHICYAAFHFSENLLLETTSAPLDEIGATFYKSGDRLSFFFQNESTLNYEHSYKNV
                     IKYVCKTFFPASNRFVYHKEPMCTRVNTWFCKFTKVDTYFLFRGVYTRGEDSEQFYTA
                     MDEAWEYKKTLAMLNSERTIFRDRAAVNFWFPKVKDMVIVPLFDGSVTSGKMKRSEVM
                     VNKDFVYTVLNHIRTYQDKALTYKNVLSFVESIRSRVIINGVTARSEWDVDKSVLQAL
                     SMTFLLQTKLAEAKDQVVLKKFQKFDDTVTNLFWKQISDAVGDLFPSIKERLISGGFV
                     KVAEQSLQIKTPDEYITFADKLVMEYKATEELQHLDISKPLERAEKYYNALSELSVLK
                     ECDEFDITQFKNLCEEKDIDPDVVAKVIVAIMKNELTLPFKNPTPEALSDALSPLPKD
                     LDMRFDLLKLSTCAPFPSVKTLDSGLLPKQSYGDERQFESQSVVSVSDFHLKSVESVK
                     MKNMSSAVYTGPLKVQQMKNYMDYLSASISATVSNLCKVLKDVYGADPESAEKSGVYD
                     VVKGKWLLLKPKDKCHAWGVAELNNGEKVIVLLEWADGFPICGDWRRVAVSSDSLIYSD
                     MGKLQTLLSCLKDGEPVPSDAKVTLVDGVPGCGKTKEILETVNFDEDLILVPGKEACK
                     MIIKRANKSGHVRATKDNVRTVDSFLMHLKPKTYNKLFIDEGLMLHTGCVNFLIALSH
                     CREAMVFGDTEQIPFINRVANFPYPKHFATLVYDHREVRRLSLRCPADVTHFMNSKYD
                     GKVLCTSDVIRSVDAEVVRGKGVFNPKSKPLKGKIITFTQSDKAELKERGYEEVSTFG
                     EVNTVHEIQGETFEDVSVVRLTPTPLELISKSSPHVLVALTRHTKSFKYYSVVLDPLV
                     KVCSDLSKVSDFILDMYKVDAGIL"
     gene            4808. .5719
                     /gene="MP"
     CDS             4808. .5719
                     /gene="MP"
                     /codon_start=1
                     /product="movement protein"
                     /protein_id="AAS87225.1"
                     /db_xref="GI:46309862"
                     /translation="MGRLRFVVLLSIFPIKTFSEPCSTMALVLRDSIKISEFINLSAS
                     EKLLPSALTAVKSVRISKVDKIISYENDTLSDIDLLKGVKLVENGYVCLAGLVVTGEW
                     NLPDNCKGGVSICLVDKRMKRANEATLGSYHTSACKKRFTFKIIPNYSVTTADALKGI
                     WQVMTNIRGVEMEKGFCPLSLEFVSICIVYLNNIKLGLREKILNVTEGGPTELTEAVV
                     DEFVEKVPMAARLKSFRSVNKKKPSNSSKFVNGKSRLNSRNKLNYENGDSDVGTSVVD
                     DIVVSNGVSDIRIDDDCESFDAQSDSY"
     gene            5722. .6198
                     /gene="CP"
     CDS             5722. .6198
                     /gene="CP"
                     /codon_start=1
                     /product="coat protein"
                     /protein_id="AAS87226.1"
                     /db_xref="GI:46309863"
                     /translation="MSYTITDPSKLAYLSSAWADPNSLINLCTNSLGNQFQTQQARTT
                     VQQQFADVWQPVPTLTSRFPAGAGYFRVYRYDPILDPLITFLMGTFDTRNRIIEVENP
                     QNPTTTETLDATRRVDDATVAIRSAINNLLNELVRGTGMYNQVSFETMSGLTWTSS"
ORIGIN

  Query Match            60.3%;  Score 45.8;  DB 14;  Length 6612;
  Best Local Similarity  76.7%;  Pred. No. 7.3e-05;
  Matches   56;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              || ||||| ||   ||| |  ||||| ||||||   |||||||| ||||||||||  ||
Db       5420 GTAACAGAAGGAGGGCCCACCGAACTTACTGAAGCAGTTGTTGATGAGTTCGTGGAGAAA 5479

Qy         61 GTACCAATGGCTG 73
              || || |||||||
Db       5480 GTTCCTATGGCTG 5492
```

```
RESULT 13
TOTMV2
LOCUS       TOTMV2                  1004 bp    RNA     linear   VRL 05-OCT-1996
DEFINITION  TMV 3' end.
ACCESSION   V01406
VERSION     V01406.1  GI:62119
KEYWORDS    coat protein.
SOURCE      Tobacco mosaic virus
  ORGANISM  Tobacco mosaic virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE   1  (bases 1 to 1004)
  AUTHORS   Guilley,H., Jonard,G., Kukla,B. and Richards,K.E.
  TITLE     Sequence of 1000 nucleotides at the 3' end of tobacco mosaic virus
            RNA
  JOURNAL   Nucleic Acids Res. 6 (4), 1287-1308 (1979)
  MEDLINE   79201054
   PUBMED   109810
COMMENT     KST TMV2.
FEATURES             Location/Qualifiers
     source          1. .1004
                     /organism="Tobacco mosaic virus"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:12242"
     CDS             321. .800
                     /note="unnamed protein product; reading frame (coat
                     protein)"
                     /codon_start=1
                     /protein_id="CAA24685.1"
                     /db_xref="GI:62120"
                     /db_xref="GOA:P03570"
                     /db_xref="Swiss-Prot:P03570"
                     /translation="MSYSITTPSQFVFLSSAWADPIELINLCTNALGNQFQTQQARTV
                     VQRQFSEVWKPSPQVTVRFPDSDFKVYRYNAVLDPLVTALLGAFDTRNRIIEVENQAN
                     PTTAETLDATRRVDDATVAIRSAINNLIVELIRGTGSYNRSSFESSSGLVWTSGPAT"
ORIGIN

  Query Match            57.9%;  Score 44;  DB 14;  Length 1004;
  Best Local Similarity  73.7%;  Pred. No. 0.00021;
  Matches   56;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              |||| ||||||   ||| || ||||| || ||| | ||||||||  |||||| |||| ||
Db         52 GTGAGAGACGGAGGGCCCATGGAACTTACAGAAGAAGTTGTTGATGAGTTCATGGAAGAT 111

Qy         61 GTACCAATGGCTGTGA 76
              || || ||| |   |
Db        112 GTCCCTATGTCAATCA 127


RESULT 14
E00088
LOCUS       E00088                  1019 bp    RNA     linear   PAT 28-SEP-1997
DEFINITION  Genomic RNA containing gene of capsid protein.
ACCESSION   E00088
VERSION     E00088.1  GI:5708407
KEYWORDS    JP 1983051894-A/2.
SOURCE      Tobacco mosaic virus
  ORGANISM  Tobacco mosaic virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE   1  (bases 1 to 1019)
  AUTHORS   Roorensu,I.P. and Merii,K.H.
  TITLE     RNA PLANT VIRUS VECTOR OR FABRICATION AND UTILIZATION OF PART
  JOURNAL   Patent: JP 1983051894-A 2 26-MAR-1983;
            NATL RES KAUNSURU OBU CANADA
COMMENT     OS   tobacco mosaic virus
            PN   JP 1983051894-A/2
            PD   26-MAR-1983
            PF   27-MAY-1982 JP 1982090482
            PR   27-MAY-1981 US 81    267539
            PI   ROORENSU II PERUCHIYAA, MERII KURISUTEIN HARASA PC
            C12N15/00,A01H1/00,C07H21/04,C12P19/34,C12P21/00//C12R1/91; CC
            strandedness: Single;
```

```
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            FH   Key             Location/Qualifiers
            FH
            FT   5'UTR           1. .320
            FT   misc_feature    61. .234
            FT                   /note='region essential for formation of  FT
                              nuclear'
            FT   misc_feature    235. .320
            FT                   /note='region for regulation' FT   CDS
                          321. .749
            FT                   /product='capsid protein'
            FT   3'UTR           750. .1022.
FEATURES             Location/Qualifiers
     source          1. .1019
                     /organism="Tobacco mosaic virus"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:12242"
ORIGIN

  Query Match             57.9%;  Score 44;  DB 6;  Length 1019;
  Best Local Similarity   73.7%;  Pred. No. 0.00021;
  Matches   56;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              |||| ||||||    ||| || |||| || ||| | |||||||| |||||| |||| ||
Db         52 GTGAGAGACGGAGGGCCCATGGAACTTACAGAAGAAGTTGTTGATGAGTTCATGGAAGAT 111

Qy         61 GTACCAATGGCTGTGA 76
              || || ||| |  | |
Db        112 GTCCCTATGTCAATCA 127


RESULT 15
D63809
LOCUS       D63809                6395 bp    RNA     linear   VRL 13-FEB-1999
DEFINITION  Tabacco mosaic virus genomic RNA for 130K protein, 180K protein,
            30K protein and coat protein, complete sequence.
ACCESSION   D63809
VERSION     D63809.1  GI:1619995
KEYWORDS    coat protein; 30K protein; 180K protein; 130K protein.
SOURCE      Tobacco mosaic virus
  ORGANISM  Tobacco mosaic virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE   1  (bases 1 to 6395)
  AUTHORS   Chen,J., Watanabe,Y., Sako,N., Ohshima,K. and Okada,Y.
  TITLE     Complete nucleotide sequence and synthesis of infectious in vitro
            transcripts from a full-length cDNA clone of a rakkyo strain of
            tobacco mosaic virus
  JOURNAL   Arch. Virol. 141 (5), 885-900 (1996)
  MEDLINE   96265021
   PUBMED   8678834
REFERENCE   2  (bases 1 to 6395)
  AUTHORS   Sako,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-AUG-1995) Nobumichi Sako, Faculty of Agriculture,
            Saga University, Laboratory of Plant Virology; 1 Honjo-machi, Saga,
            Saga 840, Japan (Tel:0952-24-5191(ex.2730), Fax:0952-22-6274)
FEATURES             Location/Qualifiers
     source          1. .6395
                     /organism="Tobacco mosaic virus"
                     /mol_type="genomic RNA"
                     /strain="Rakkyo"
                     /db_xref="taxon:12242"
     CDS             69. .4919
                     /note="position 3417. .3419 is readthrough codon"
                     /codon_start=1
                     /trans_except=(pos:3417. .3419,aa:OTHER)
                     /product="180K protein"
                     /protein_id="BAA09876.1"
                     /db_xref="GI:1619996"
                     /translation="MAYTQTATTSALLDTVRGNNSLVNDLAKRRLYDTAVDEFNARDR
                     RPKVNFSKVISEEQTLIATRAYPEFQITFYNTQNAVHSLAGGLRSLELEYLMMQIPYG
                     SLTYDIGGNFASHLFKGRAYVHCCMPNLDVRDIMRHEGQKDSIELYLSRLDRGGKTVP
                     NFQKEAFDRYAEFPEDAVCHNTFQTCEHQPMQQSGKVYAIALHSIYDIPADEFGAALL
                     RKNVHTCYAAFHFSENLLLEDSYVNLDEINACFSRDGDKLTFSFASESTLNYCHSFSN
                     ILKYVCKTYFPASNREVYMKEFLVTRVNTWFCKFSRIDTFLLYKGVAHKSVDSEQFYT
                     AMEDAWHYKKTLAMCNSERILLEDSSSVNYWFPKMRDMVIIPLFDISLETSKRSRKEV
                     LVSKDFVFTVLNHIRTYQAKALTYANVLSFVESIRSRVIINGVTARSEWDVDKSLLQS
                     LSMTFFLHTKLAVLKDDLLISKFSLGSKTVCQHVWDEISLAFGNAFPSVKERLLNKKL
                     IRAAGDALEIKVPDLYITFHDRLVAEYKSSVDMPALDIRKRMEETEVMYNALSELSVL
                     RESDKFDVDVFSQMCKSLEVDPMTAAKVIVAVMSNESGLTLTFERPTEANVAQALQDQ
                     EKASEGALVVTSREVEEPSMKGSMARGELQLAGFAGDHPESSYSRNEEIESLEQFHMA
                     TADSLIRKQMSSIVYTGPIKVQQMKNFIDSLVASLSAAVSNLVKILKDTAAIDLETRQ
                     KFGVLDVASRKWLIKPTAKSHAWGVVETHARKYHVALLEYDEQGIVTCDDWRRVAVSS
                     ESVVYSDMAKLRTLRRLLRDGEPHVSNAKVVLVDGVPGCGKTKEILSRVNFDEDLILV
                     PGKQAAEMIRRRANSSGIIVATKDNVRTVDSFMMNFGKTTRCQFKRLFIDEGLMLHTG
                     CVNFLVAMSLCDVAYVYGDTQQIPYINRVSGFPYPAHFSKLEVDEVETRRTTLRCPAD
                     VTHYLNRRYEGFVVSTSSVKKSVSQEMVSGAAVINPISKPLHGKILTFTQSDKEALLS
                     RGYSEVHTVHEVQGETYSDVSLVRLTPTPISIIAGDSPHVLVALSRHTCSLKYYTVVM
                     DPLVSIIRDLEKLSSYLLDMYKVDAGIQXQLQIDSVFKGSNLFVAAPKTGDISDMQFY
                     YDKCLPGNSTMMNNFDAVTMRLTDISLNVKDCILDMSKSVAAPKDQIKPLIPMVRTAA
                     EMPRQTGLLENLVAMIKRNFNAPELSGIIDIENTASLVVDKFFDSYLLKEKRKPNKNV
                     SLFSRESLNRWLEKQERVTIGQLADFDFVDLPAVDQYRHMIKAQPKQKLDTSIQTEYP
                     ALQTIVYHSKKINAIFGPLFSELTRQLLDSVDSSRFLFFTRKTPAQIEDFFGDLDSHV
                     PMDVLELDISKYDKSQNEFHCAVEYEIWRRLGFEDFLGEVWKQGHRKTTLKDYTAGIK
                     TCIWYQRKSGDVTTFIGNTVIIAACLASMLRMEKIIKGAFCGDDSLLYFPKGCEFPDI
                     QHSVNLMWNFEAKLFKKQYGYFCGRYIIHHDRGCIVYYDPLKLISKLGAKHIKDWEHL
                     EEFRRSLCDVAVSLNNCAYYTQLDDAVWEVHKTAPPGSFVYKSLVKYLSDKVLFRSLF
                     INGSSC"
     CDS             69. .3419
                     /codon_start=1
                     /product="130K protein"
                     /protein_id="BAA09877.1"
                     /db_xref="GI:1619997"
                     /translation="MAYTQTATTSALLDTVRGNNSLVNDLAKRRLYDTAVDEFNARDR
                     RPKVNFSKVISEEQTLIATRAYPEFQITFYNTQNAVHSLAGGLRSLELEYLMMQIPYG
                     SLTYDIGGNFASHLFKGRAYVHCCMPNLDVRDIMRHEGQKDSIELYLSRLDRGGKTVP
                     NFQKEAFDRYAEFPEDAVCHNTFQTCEHQPMQQSGKVYAIALHSIYDIPADEFGAALL
                     RKNVHTCYAAFHFSENLLLEDSYVNLDEINACFSRDGDKLTFSFASESTLNYCHSFSN
                     ILKYVCKTYFPASNREVYMKEFLVTRVNTWFCKFSRIDTFLLYKGVAHKSVDSEQFYT
                     AMEDAWHYKKTLAMCNSERILLEDSSSVNYWFPKMRDMVIIPLFDISLETSKRSRKEV
                     LVSKDFVFTVLNHIRTYQAKALTYANVLSFVESIRSRVIINGVTARSEWDVDKSLLQS
                     LSMTFFLHTKLAVLKDDLLISKFSLGSKTVCQHVWDEISLAFGNAFPSVKERLLNKKL
                     IRAAGDALEIKVPDLYITFHDRLVAEYKSSVDMPALDIRKRMEETEVMYNALSELSVL
                     RESDKFDVDVFSQMCKSLEVDPMTAAKVIVAVMSNESGLTLTFERPTEANVAQALQDQ
                     EKASEGALVVTSREVEEPSMKGSMARGELQLAGFAGDHPESSYSRNEEIESLEQFHMA
                     TADSLIRKQMSSIVYTGPIKVQQMKNFIDSLVASLSAAVSNLVKILKDTAAIDLETRQ
                     KFGVLDVASRKWLIKPTAKSHAWGVVETHARKYHVALLEYDEQGIVTCDDWRRVAVSS
                     ESVVYSDMAKLRTLRRLLRDGEPHVSNAKVVLVDGVPGCGKTKEILSRVNFDEDLILV
                     PGKQAAEMIRRRANSSGIIVATKDNVRTVDSFMMNFGKTTRCQFKRLFIDEGLMLHTG
                     CVNFLVAMSLCDVAYVYGDTQQIPYINRVSGFPYPAHFSKLEVDEVETRRTTLRCPAD
                     VTHYLNRRYEGFVVSTSSVKKSVSQEMVSGAAVINPISKPLHGKILTFTQSDKEALLS
                     RGYSEVHTVHEVQGETYSDVSLVRLTPTPISIIAGDSPHVLVALSRHTCSLKYYTVVM
                     DPLVSIIRDLEKLSSYLLDMYKVDAGIQ"
     CDS             4903. .5709
                     /codon_start=1
                     /product="30K protein"
                     /protein_id="BAA09878.1"
                     /db_xref="GI:1619998"
                     /translation="MALVVKGKVNINEFIDLTKMEKILPSMFTPVKSVMCSKVDKIMV
                     HENESLSEVNLLKGVKLIDSGYVCLAGLVVTGEWNLPDNCRGGVSVCLVDKRMERADE
                     ATLGSYYTAAAKKRFQFKVVPNYAITTQDAMKNVWQVLVNIRNVKMSAGFCPLSLEFV
                     SVCIVYRNNIKIGLREKITNVRDGGPMELTEEVVDEFMEDVPMSIRLAKFRSRTGKKS
                     VVPKGNFSSRDRSQPNKNYGNAKDFGGMSFKKNNLIDDGSETSVAESDSF"
     CDS             5712. .6191
                     /codon_start=1
                     /product="coat protein"
                     /protein_id="BAA09879.1"
                     /db_xref="GI:1619999"
                     /translation="MSYNINTPSQFVFLSSAWADPIELINLCTNALGNQFQTQQARTV
                     VQRQFSEVWKPSPQVTVRFPDSDFKVYRFNAVLDPLVTALLGAFDTRNRIIEVENQAN
                     PSTAETLDATRRVDDATVAIRSAINNLIVELTRGTGSYNRSSFESSSGLVWTSSPAT"

ORIGIN
```

Query Match 57.9%; Score 44; DB 14; Length 6395;
Best Local Similarity 73.7%; Pred. No. 0.00031;
Matches 56; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```
Qy          1 GTGACAGACGGCTCGCCAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
              |||| ||||||     || || ||||| || ||| | |||||||| |||||| |||| ||
Db       5443 GTGAGAGACGGGGGTCCCATGGAACTTACAGAAGAAGTTGTTGATGAGTTCATGGAAGAT 5502

Qy         61 GTACCAATGGCTGTGA 76
              ||||| ||| |  | |
Db       5503 GTACCTATGTCAATTA 5518
```

Search completed: December 2, 2004, 08:20:47
Job time : 1637 secs